Application of the Analysis of Variance and Covariance Method to Educational Problems

BULLETIN No. 11

OF THE

DEPARTMENT OF EDUCATIONAL RESEARCH

BY

ROBERT W. B. JACKSON, B.A. (Alta.) Ph.D. (London)

The preparation of this Bulletin was aided by a grant from the Canadian Council for Educational Research

DEPARTMENT OF EDUCATIONAL RESEARCH
UNIVERSITY OF TORONTO
371 BLOOR STREET WEST
TORONTO 5

Application of the Analysis of Variance and Covariance Method to Educational Problems

BULLETIN No. 11

OF THE

DEPARTMENT OF EDUCATIONAL RESEARCH

BY

ROBERT W. B. JACKSON, B.A. (Alta.) Ph.D. (London)

The preparation of this Bulletin was aided by a grant from the Canadian Council for Educational Research

PRICE $1.00

DEPARTMENT OF EDUCATIONAL RESEARCH
UNIVERSITY OF TORONTO
371 BLOOR STREET WEST
TORONTO 5

PRINTED BY
THE UNIVERSITY OF TORONTO PRESS
1940

REPRODUCED BY
HILL LITHOGRAPHING, TORONTO, FEBRUARY 1948

FOREWORD

The author of this Bulletin, Dr. R. W. B. Jackson, after graduating with first-class honours in Mathematics from the University of Alberta, continued his studies in the University of London, England. There he worked for five years under R. A. Fisher, Egon Pearson and others, and became fully seized of the value of the new statistical techniques they had developed. Consequently, when he joined our staff in the spring of 1939, I suggested that he should prepare a bulletin which would make available to research students in education the statistical techniques associated with Fisher's name. What I had in mind was a publication that would do for educators what Snedecor's *Calculation and Interpretation of Analysis of Variance and Covariance* had done for agriculturalists. This latter work, using data from biological experiments, had expounded Fisher's techniques of analysis of variance and covariance in the solution of agricultural problems. Could not the data from educational tests and research, which we had in abundance in our Department, be put to similar use? Research students in education should be shown by practical demonstration when and how to use the variance and covariance method of Fisher.

As soon as the Canadian Council for Educational Research was established, I applied, on Dr. Jackson's behalf, for a modest grant to aid this particular project. The application was well received and a grant-in-aid was made. For the grant, grateful acknowledgement is hereby made.

In my first enthusiasm for the project, I envisaged a bulletin that would be understandable by a person with a modicum of mathematical and statistical knowledge, but I am now obliged to confess that Fisher's contributions to statistics cannot be explained in mathematical words of one syllable. Nevertheless, Dr. Jackson has succeeded in this bulletin in showing how the illustrative problems are worked out step by step, and what deductions can rightfully be made from the results. Practically all the types of cases are given in which the technique of analysis of variance can be used on educational data. In the last chapter, dealing with analysis of covariance, only data exhibiting linearity of regression are considered, but similar methods may be used in other cases.

I feel that this short, practical work of Dr. Jackson's should find a warm welcome among research workers in education. In any case, if it extends the knowledge of a valuable statistical technique among them, the bulletin will have served its purpose.

PETER SANDIFORD, *Director*.

University of Toronto
January 1940.

ACKNOWLEDGEMENTS

I wish to acknowledge my indebtedness to the following authors for their kind permission to reproduce the tables given in the appendices:

(1) Professor R. A. Fisher, and his publishers Oliver S. Boyd, Edinburgh, for permission to reproduce the tables giving the 5% and 1% points of the distribution of z.

(2) Professor George W. Snedecor, and his publishers the Collegiate Press, for permission to reproduce the tables giving the 5% and 1% points for the distribution of F.

(3) Professor Egon S. Pearson for permission to reproduce the tables giving the 5% and 1% limits for the distribution of L_1.

My thanks are due also to Professors A. E. Brandt and Palmer O. Johnson for reading and criticizing the manuscript and for their many helpful suggestions; and to members of our Department, in particular to Miss K. Hobday and Miss M. Graham, for their assistance in this work.

R. W. B. JACKSON.

CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION AND OUTLINE OF THE THEORY UNDERLYING THE STATISTICAL TESTS USED IN THE ANALYSIS OF VARIANCE AND COVARIANCE

The statistical method known as the Analysis of Variance and Covariance is one of the most useful and powerful that has been developed in recent years. It was first suggested and used by Professor R. A. Fisher in England (in 1923) and is at present one of the standard methods of procedure for analyzing the results of agricultural and biological experiments. Its usefulness, however, is not limited to these fields, as it is a method which seems to admit of general application, and it should prove to be of particular value in the field of education.[1] It is the purpose of the present Bulletin to outline the general principles underlying the method and to show, by examples, how it may be applied to particular problems in the educational field.

The method consists, as the name implies, in the breaking-up of the total variance (or covariance) into independent parts which may

[1]The method has been used in a few cases but it is certainly not as yet in general use. Two interesting articles have been published since this bulletin was written: the first by Paul L. Dressel, on "The Effect of High School on College Grades", and the second by Jack W. Dunlap, "Applications of Analysis of Variance to Educational Problems." Dunlap's article gives a brief discussion of the Analysis of Variance method and a summary of the results given in articles in which this method has been applied to educational problems. A list of these articles is given below; it will be noted that the majority of them have been published in the last two years.

References;

(1) Dressel, Paul L. "The Effect of High School on College Grades". *Journal of Educational Psychology*, XXX (1939), pp. 612-617.

(2) Dunlap, Jack W. *Race Differences in the Organization of Numerical and Verbal Abilities*. Archives of Psychology, 1931, No. 124. Pp. 72.

(3) Dunlap, Jack W. "Applications of Analysis of Variance to Educational Problems". *Journal of Educational Research*, XXXIII (1940), pp. 434-442.

(4) Lev, Joseph. "Evaluation of Test Items by the Method of Analysis of Variance". *Journal of Educational Psychology*, XXIX (1938), pp. 623-630.

(5) Rubin-Rabson, Grace. "Studies in the Psychology of Memorizing Piano Music. I. A Comparison of the Unilateral and the Coordinated Approaches". *Journal of Educational Psychology*, XXX (1939), pp. 321-345.

be ascribed to certain known factors or components. It must not be confused with the factor-analysis methods, based on a similar principle, now used in psychology and education. In the method considered here, the factors or components are known and specified beforehand, and the experiments must be carefully planned and arranged in order that the influence of these factors may be isolated and measured. We obtain, at the same time, an estimate of the experimental error free from the effect of these known and measurable factors. The use of the degrees of freedom,[2] also introduced by Fisher, makes the method equally valid for large and small samples.

The tests of significance of the various factors are based on the comparison of two estimates of variance: the variance ascribable to the factor under consideration and to the experimental error. The tests all reduce to the single one of determining whether one estimate of variance, with n_1 degrees of freedom, is significantly greater than a second independent estimate of the same variance with n_2 degrees of freedom. The problem involved in making this test was solved by Fisher(**1**), and the tables of z[3] which he has prepared and published, or the tables of F[4] published by Snedecor(**11**), enable us to make the test in the particular cases in which we are interested. The use of these tables will be explained and illustrated in the problems which follow.

The term "variance" as used in this method refers to the square of the standard deviation, i.e.

$$V=\sigma^2 \tag{1}$$

where V represents the variance and σ the standard deviation. The "covariance" is related to the correlation coefficient and the standard deviations of two related variables. This relation may be expressed as

$$C_v=\rho_{12}\,\sigma_1\,\sigma_2 \tag{2}$$

where C_v represents the covariance, σ_1 and σ_2 the standard deviations of the first and second variables, respectively, and ρ_{12} represents the correlation coefficient. In the analysis of variance, therefore, we consider the variation of only one quantity, while in the analysis of covariance we consider the associated variation, or the covariation,

[2]The term "degrees of freedom" refers to the number used as divisor in our estimate of the variance. The bias of our estimate in small samples is compensated, on the average, by using the number of degrees of freedom instead of the number of observations. (See equation (13).)

[3]Tables VI and VII, pp. 250-253. (See Appendix A.)

[4]Table 10.3, pages 184-187. (See Appendix B.)

of two quantities. It will readily be seen that the analysis of covariance is a natural extension of the general analysis of variance method.

In estimating the variance, V, we first calculate the "sum of squares", as it is known, which is simply an abbreviated phrase denoting the sum of squares of the deviations of a set of observed values from their mean. It is, actually, the total sum of squares rather than the variance which is broken up into parts in the analysis. The additive properties, to be discussed later, are common to both the sum of squares and the degrees of freedom, and not, strictly speaking, to the estimates of the variance. The general term "mean square" is used to denote the various estimates of variance which are calculated; it is, as the name suggests, the quotient of a sum of squares and the appropriate number of degrees of freedom. If we denote by X any value of a variable, by $\overline{X}$ the mean value of X as calculated from a sample, by $\sum$ summation or addition, and by f the corresponding number of degrees of freedom, then the sum of squares may be written[5]

$$\sum(X-\overline{X})^2 \tag{3}$$

and the mean square

$$\frac{1}{f}\sum(X-\overline{X})^2 \tag{4}$$

The statistical theory underlying the method and the tests of significance will not be re-developed here as it would be of little interest to the practical research worker. It is suggested that readers who are interested in this side of the problem should refer to the texts of Fisher, and Snedecor (**1**, **10**, **11**) and in particular to interesting papers by Kolodziejczyk (**4**) and Johnson and Neyman (**3**) on this general problem. At the same time, however, it is felt that even the practical research worker will benefit from a study of the general principles underlying the theory since this knowledge will aid him in deciding whether or not the method is applicable to his particular problem, and will also help him in learning how to use it. In addition, it will enable him to attack new problems with more confidence and guide him in planning experiments designed to solve them. The next few pages, therefore, will be devoted to a general outline of the theory and of the procedure to be followed in attacking such problems. This section may be omitted by the non-mathematical reader as many of the problems encountered will be sufficiently similar to those

[5]In the notation of Fisher, this sum of squares is written $S(X-\overline{X})^2$, and in the notation of Snedecor it is written $\Sigma X^2-(\Sigma X)M_X$ or $\Sigma X^2-(\Sigma X)^2/n$.

discussed later that they may be solved without a detailed study of the theory.

In solving a statistical problem, the statistician must first set up a statistical model to which the problem conforms with sufficient accuracy for results of practical value to be deduced. The model to which problems seem to conform in this particular type of analysis is one in which a particular observation, X_i, is assumed to consist of a random element, z_i, plus a linear function, μ_i, of a specified number of constants or parameters[6]. If we denote the parameters by $\theta_1, \theta_2, \ldots \ldots, \theta_q$, then we may express these assumptions in mathematical form. Thus

$$\mu_i = C_{1i}\,\theta_1 + C_{2i}\,\theta_2 + \ldots + C_{qi}\,\theta_q \tag{5}$$

where the quantities $C_{1i}, C_{2i}, \ldots, C_{qi}$ are known values—generally $0, +1$, or -1—determined by the nature of the problem under consideration. From (5), we may write

$$\begin{aligned} X_i &= \mu_i + z_i \qquad (i = 1, 2, \ldots\ldots, N) \\ &= C_{1i}\,\theta_1 + \ldots\ldots + C_{qi}\,\theta_q + z_i \end{aligned} \tag{6}$$

It is assumed, further, that these random elements, in successive observations, are independent of each other and of the values of the parameters, and are normally distributed about zero, with the same standard deviation σ in all cases, i.e.

$$p(z_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-z_i^2/2\sigma^2} \tag{7}$$

where $p(z_i)$ denotes the probability distribution of z_i. From equation (6), it follows that

$$p(z_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(X_i - \mu_i)^2}{2\sigma^2}} \tag{8}$$

and the probability distribution of all the z's will be

$$p(z_1, \ldots, z_n) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N e^{\frac{-\sum_{i=1}^{N} (X_i - \mu_i)^2}{2\sigma^2}} \tag{9}$$

In equations (7), (8), (9), σ is the standard deviation of z_i in the population from which we are sampling. It is not necessary to know

[6]Parameters are constants describing the properties of the population from which we are sampling.

its precise value as the tests of significance are independent of the particular value of σ. In educational problems the estimate of σ is very useful, but this is a different question, statistically speaking.

The value of σ, as we see from the above assumptions, is supposed to be the same whatever the value of the parameters $\theta_1, \theta_2, \ldots\ldots, \theta_q$. Methods have been developed which may be used when this assumption is not satisfied but these will not be considered here. Experience shows that most of the problems conform to the above model with sufficient accuracy for our purpose.

Up to this point, we have defined the model only in a general form; these fundamental assumptions and conditions are common to all the problems. In a given situation, i.e. when we are faced with some practical problem, the model will have to be adjusted to meet the specific conditions which govern it. Various hypotheses regarding the factors which contribute to the variation in X are conceivable, and we must first determine the minimum number, say s, of parameters θ that are needed to define exactly the set of admissible hypotheses. The precise values of the parameters are unknown, but the hypotheses will specify either certain values of the parameters (zero, for example) or their relations to one another. A particular hypothesis will be defined by one or more relations, in general say r relations where $r<s$, between the s parameters $\theta_1, \theta_2, \ldots\ldots, \theta_s$. The data collected, consisting of N observations $X_1, X_2, \ldots\ldots, X_N$, may then be used to test this hypothesis.

Before discussing the method of developing such a test, let us digress for a moment to consider more exactly what we mean by the phrase "testing a hypothesis". The set of admissible hypotheses consists of all possible hypotheses regarding the unknown values of the parameters or their relations to one another which are relevant to our problem, or, as is more frequently the case, which we admit to be relevant. The particular hypothesis which we wish to test will be defined by the r relations between the parameters. Within the set of admissible hypotheses there will be others which are alternative to this one, i.e. which are contradictory to it. Let us denote by H_0 the hypothesis to be tested[7], by H^1 some other hypothesis contradictory to H_0, by W the sample space, and by w any region in this sample space. The coordinates of the sample space, W, are the

[7]In the examples which follow there is frequently more than one hypothesis to be tested; these have been denoted by H_0, H_1, H_2, H_3, ..., etc. Since the problems considered in the different examples are not the same, the symbols, H_0 for example, may refer to different hypotheses in different examples. In each case, of course, the hypothesis to be tested has been defined and discussed.

possible results of experiment; and the observed results in any particular experiment will define a point, E, called the sample point, in W. Denote by $p\{E\epsilon w\}$ the probability that E, as determined by some experimental results, will fall within the region w, and by $p\{E\epsilon w/H\}$ the probability of E falling within w as determined by some hypothesis, H.

Any test of a statistical hypothesis, H_0, reduces to the rule of rejecting H_0 when the sample point, E, falls within some specified region, w_0, called critical, and of accepting H_0 (or at least not rejecting it) in all other cases. There are two kinds of errors which may arise in testing statistical hypotheses: (1) we may reject the hypothesis tested, H_0, when it is true, and (2) we may fail to reject H_0 when it is false, i.e. when some alternative hypothesis, H^1, is, in fact, true. The probability of the first kind of error is $p\{E\epsilon w_0/H_0\}$ and the probability of the second kind of error is $p\{E\epsilon(W\text{-}w_0)/H^1\}$ where $W-w_0$ denotes the part of W outside of w_0.

The theory of testing statistical hypotheses as developed by Neyman and Pearson (**6**, **7**, **8**, **9**) is based on the simple concept of arranging the test, i.e. of choosing the critical region w_0, so as to minimize the probability of errors. The critical regions used in the tests of the statistical hypotheses in this bulletin appear to be the best possible in the sense that, (1) the probability of errors of the first kind is controlled at a fixed level, and (2) the probability of errors of the second kind seems to be reduced to as low a level as possible for the class of hypotheses in which we are interested.

Our problem of developing a test of a particular hypothesis reduces, therefore, to that of determining a critical region w_0 possessing the above properties. The appropriate region in our case is determined by a comparison of sums of squares; these are obtained as follows:

Denote by

$$\chi^2 = \sum_{i=1}^{N} \{X_i - m_i\}^2 \tag{10}$$

where

$$m_i = C_{1i}q_1 + C_{2i}q_2 + \;\ldots\ldots\; + C_{si}q_s \tag{11}$$

is similar to μ_i of (5) except that the unknown parameters θ are replaced by continuous variables q.

Step 1.

Minimize χ^2 with regard to the s unknown q's and denote the resulting absolute minimum by χ^2_a. This consists in obtaining the

derivatives of χ^2 with regard to all the q's, equating these to zero, and solving the resulting equations for the values of q_i, say q_i^0, which minimize χ^2. We obtain the required minimum value of χ^2, χ_a^2, by substituting q_i^0 in (10).

Step 2.

Minimize χ^2 subject to the condition that the s variables q_i satisfy the r relations specified by the hypothesis tested. Proceeding as in Step 1, we obtain the required relative minimum value of χ^2, denoted by χ_r^2.

Step 3.

Calculate the quantity

$$z=\tfrac{1}{2}\log_e\left[\frac{\chi_r^2-\chi_a^2}{f_1}\Big/\frac{\chi_a^2}{f_2}\right] \tag{12}$$

where f_1 and f_2 denote degrees of freedom, defined as follows:

$$\begin{aligned} f_1 &= r \\ f_2 &= N-s \end{aligned} \tag{13}$$

Step 4.

Using this quantity[3], refer to Fisher's tables of z (these are entered with the above degrees of freedom which Fisher denotes by n_1 and n_2, respectively) and

(a) reject the hypothesis tested, H_0, if the calculated value of z is greater than the 5% (or 1%) point given in the tables,

or (b) accept the hypothesis tested, H_0, if the calculated value of z is less than the 5% (or 1%) point given in the tables.

The question of whether to use the 5%, 1%, or 0.1% point as the boundary of the critical region is a personal one, and depends on what probability of the first kind of error we consider as permissible. The general practice seems to be to reject the hypothesis tested if z is greater than, i.e. lies beyond, the 1% point, to remain in doubt if it lies between the 5% and 1% points, and to accept the hypothesis tested if z is less than the 5% point. It is impossible to give a definite ruling because in many cases the particular level to be used will be determined by the nature of the problem under consideration.

[3]If Snedecor's tables of F are to be used, we calculate $F=\dfrac{\chi_r^2-\chi_a^2}{f_1}\Big/\dfrac{\chi_a^2}{f_2}$ instead of z. The procedure explained below is applicable to both cases.

Notes on the use of Fisher's tables of z.

(1) *Example* (Artificial)

Let us assume that our experimental results give us the following values for the test of a statistical hypothesis, H_0.

$$\begin{aligned} f_1 &= r = 6 \\ f_2 &= N - s = 60 \\ z &= 0.8 \end{aligned} \tag{14}$$

We proceed as follows: from Table VI, pages 250, 251 (1), or Appendix A, we find from the column headed $n_1=6$ and the row headed $n_2=60$, that the 5% point of the distribution of z is 0.4064; from Table VI, pages 252, 253, (1) we find from the column headed $n_1=6$ and the row headed $n_2=60$, that the 1% point of the distribution of z is 0.5687. Our calculated value of z, $z=0.8$, is greater than the 1% point so our sample point, E, falls within the critical region and we reject the hypothesis to be tested, H_0.

(2) The tables are to be entered with degrees of freedom n_1 corresponding to the larger mean square. In the event of

$$\frac{\chi_r^2-\chi_a^2}{r} < \frac{\chi_a^2}{N-s} \tag{15}$$

we calculate

$$z^1 = \tfrac{1}{2}\log_e\left[\frac{\chi_a^2}{N-s} \Big/ \frac{\chi_r^2-\chi_a^2}{r}\right] \tag{16}$$

and enter the tables with degrees of freedom,

$$\begin{aligned} n_1 &= N - s \\ n_2 &= r \end{aligned} \tag{17}$$

Otherwise the procedure is the same as that outlined above in the first example. The interpretation of the results, however, is sometimes difficult in the case of the rejection of the hypothesis to be tested, H_0. The clue to the interpretation in these cases can generally be found by a close examination of the nature of the hypotheses alternative to the one tested.

(3) The 5% and 1% points of the distribution of z for values of n_1 and n_2 not appearing in the tables may be obtained by simple linear interpolation. The following examples illustrate the procedure to be followed in such cases:

(a) Assume that we have $n_1=16$ and $n_2=60$ and wish to find the 5% point of the distribution of z; the nearest tabled values are for

$n_1=12$, $n_1=24$ and $n_2=60$. From Table VI, pages 250 and 251, or Appendix A, we find

n_2 \ n_1	12	24
60	0.3255	0.2654

To find the value of the 5% point corresponding to $n_1=16$, we interpolate in the above table with values of $\frac{24}{n_1}$ instead of n_1, i.e.

$\frac{24}{n_1}$	2	*1.5*	1
n_2 \ n_1	12	*16*	24
60	0.3255	*0.295*	0.2654

to obtain the desired value, 0.295.

(b) Assume that we have $n_1=20$ and $n_2=120$ and wish to find the 5% point of the distribution of z; the nearest tabled values are for $n_1=12$, $n_1=24$, $n_2=60$ and $n_2=\infty$. From Table VI, pages 250 and 251, or Appendix A, we find

n_2 \ n_1	12	24
60	0.3255	0.2654
∞	0.2804	0.2085

In this case we must perform a double interpolation, in either one of two ways, so we take $\frac{24}{n_1}$ and $\frac{60}{n_2}$ and proceed as before. We obtain

$\frac{60}{n_2}$	n_2 \ n_1	$\frac{24}{n_1}$: 2.0	1.2	1.0
		12	20	24
1.0	60	0.3255	*0.277*	0.2654
0.5	*120*	*0.303*	*0.25*	*0.237*
0	∞	0.2804	*0.223*	0.2085

The value so obtained, 0.25, is accurate to the second decimal place, which is sufficient for our purpose, as an approximate value of the probability of errors of the first kind is all we require. If we find, for example, that the value of z calculated from experimental results falls nearly at the 1% point, we know that in repeated sampling we should obtain a value as great or greater than this by chance alone in about 1% of such cases. A knowledge of the exact proportion, or the exact value of the 1% point correct to 4 decimal places, is not necessary here as we should be inclined to reject the hypothesis in any case.

The remainder of this bulletin is devoted to a study of the theory, etc., as applied to particular problems. As this method has seldom been used by others in the field of education, most of these problems are ones in which we have been particularly interested. It is hoped that enough of the field will be covered to enable other workers to apply similar methods in the solution of their own problems.

CHAPTER II

ANALYSIS OF VARIANCE

In this chapter we shall discuss the application of the Analysis of Variance method to certain educational problems and to the interpretation of the results. The examples considered are simple, but they will illustrate the use of the method. More complex problems will occur, of course, but it is suggested that the solution of these will not be difficult if the fundamental principles outlined in the first chapter and applied here are clearly understood.

It is convenient to divide the examples of this chapter into two parts:

(1) examples in which there are equal numbers of observations in the classes, and

(2) examples in which there are unequal numbers of observations in the classes.

The general method to be used is the same for both types, but the calculation of the results is simpler for the first.

PART 1

Equal Numbers of Observations in the Classes

***Example* 1.** *Resemblance of Identical Twins in Intelligence.*

This example is based on data collected by Wingfield (14) and refers to the problem of measuring the resemblance of identical twins in general intelligence. The data are given in Table I; columns 2 and 3 give the composite mental age scores of 31 identical (girl) twin pairs.

The only factors which enter here are, (1) the mental age of the groups tested, (2) the mental age of each twin pair, and (3) the difference between the mental ages of the members of each twin pair. There is one other factor, the error involved in the measurement of mental age, but it cannot be considered at present as the data are not arranged to give us a separate measure of the influence of this factor. This particular problem will be considered later (see Example 5); it is suffi-

TABLE I

MENTAL AGES OF IDENTICAL GIRL TWINS

Twin Pair (t)	Mental Age		Sum	Diff.
	X_{1t}	X_{2t}	$X_{1t}+X_{2t}$	$\lvert X_{1t}-X_{2t}\rvert$
1	178	174	352	4
2	111	117	228	6
3	102	123	225	21
4	119	124	243	5
5	137	130	267	7
6	132	135	267	3
7	136	141	277	5
8	126	123	249	3
9	188	159	347	29
10	136	140	276	4
11	137	155	292	18
12	124	118	242	6
13	99	104	203	5
14	140	137	277	3
15	166	180	346	14
16	112	121	233	9
17	117	130	247	13
18	136	148	284	12
19	143	149	292	6
20	121	122	243	1
21	184	187	371	3
22	153	167	320	14
23	135	126	261	9
24	218	191	409	27
25	156	162	318	6
26	145	121	266	24
27	165	158	323	7
28	167	186	353	19
29	127	117	244	10
30	137	142	279	5
31	143	154	297	11
Sum	4,390	4,441	8,831	—
Sum of Squares	642,732	653,179	2,587,051	4,771

cient to note here that part of the difference between the mental ages of the members of each twin pair may be due to this factor.

We may express the relationship between the factors to be considered as follows: Denote by

$$X_{it} = A + C_t + z_{it} \tag{18}$$

the mental age of the i-th member of the t-th pair of twins, where $i=1, 2; t=1, 2, \ldots, n$. A is a measure of the common mental age of the group of girls tested, and is defined as the arithmetic mean of the mental ages for all individuals; C_t is a measure of the mental age of the t-th twin pair; z_{it} is a measure of the differences between the mental ages of each twin pair. Since A is considered as the common mental age of the group tested and defined as the mean mental age for all individuals, it is necessary that

$$\sum_t C_t = 0 \tag{19}$$

The hypothesis we wish to test is

$$H_0:\ C_t = 0\ ;\ t=1, 2, \ldots, n-1 \tag{20}$$

i.e. the hypothesis that the mental age of an individual is independent of the particular twin pair to which she belongs. This is equivalent, as Fisher has shown (1, section 40, page 228), to the hypothesis that the correlation between twins (intraclass correlation) is zero.

Following the general method outlined in the first section, we first write

$$\chi^2 = \sum_i \sum_t (X_{it} - A - C_t)^2 \tag{21}$$

Minimizing χ^2 with regard to A and C_t, we obtain

$$\overset{*}{A} = \frac{1}{2n} \sum_i \sum_t X_{it} = \bar{X}_{..} \tag{22}$$

$$\overset{*}{C}_t = \tfrac{1}{2} \sum_i X_{it} - \bar{X}_{..} = X_{.t} - X_{..} \tag{23}$$

Substituting these values in (21) to obtain the absolute minimum value of χ^2, we have

$$\chi_a^2 = \sum_i \sum_t (X_{it} - X_{.t})^2 \tag{24}$$

If the hypothesis to be tested, H_0: $C_t = 0$, is true, then (21) becomes

$$\chi^2 = \sum_i \sum_t (X_{it} - A)^2 \tag{25}$$

Minimizing with regard to A, and substituting the obtained value $A = X_{..}$ in (25) we obtain the relative minimum

$$\chi_r^2 = \sum_i \sum_t (X_{it} - X_{..})^2 \tag{26}$$

We may rewrite (26) in the form

$$\chi_r^2=\sum_i\sum_t(\overline{X}_{\cdot t}-\overline{X}_{\cdot\cdot})^2+\sum_i\sum_t(X_{it}-\overline{X}_{\cdot t})^2$$
$$=\chi_b^2+\chi_a^2 \qquad , \text{ say.} \tag{27}$$

The degrees of freedom are

$$\left.\begin{aligned} f_1&=r=n-1\\ f_2&=N-s=2n-n=n\end{aligned}\right\} \tag{28}$$

For purposes of calculation it is simpler to write χ_a^2 and χ_b^2 in the form

$$\chi_a^2=\tfrac{1}{2}\sum_t(X_{1t}-X_{2t})^2 \tag{29}$$

$$\chi_b^2=\tfrac{1}{2}\left[\sum_t(X_{1t}+X_{2t})^2-\frac{(\sum_i\sum_t X_{it})^2}{n}\right] \tag{30}$$

We may check the calculations by obtaining

$$\chi_r^2=\sum_i\sum_t X_{it}^2-\frac{(\sum_i\sum_t X_{it})^2}{2n} \tag{31}$$

separately, and using the identity

$$\chi_r^2=\chi_a^2+\chi_b^2 \tag{32}$$

The easiest way to calculate the necessary values is that shown in Table I; we form the sum and difference for each pair of values and then calculate the sum and sum of squares for each column, except the last where only the sum of squares is required. This method has the added advantage of enabling us to check the calculations at each stage. From the last two rows of Table I, we find

$$\sum_t(X_{1t}-X_{2t})^2=4{,}771$$
$$\sum_t(X_{1t}+X_{2t})^2=2{,}587{,}051$$
$$\sum_i\sum_t X_{it}=8{,}831$$
$$\sum_i\sum_t X_{it}^2=1{,}295{,}911$$

Substituting these values in (29), (30), (31), we have

$$\chi_b^2=35{,}677.7419$$
$$\chi_a^2=2{,}385.5000$$
$$\chi_r^2=38{,}063.2419$$

It is convenient, and also customary, to place all these values in one table. This has been done in Table II, which also shows the notation generally used in the analysis of variance; the column headed "d.f." refers to the degrees of freedom associated with the sum of squares given in the third column.

TABLE II

ANALYSIS OF VARIANCE OF MENTAL AGES OF IDENTICAL GIRL TWINS

Variance	d.f.	Sum of Squares	Mean Square
Between Pairs	30	35,677.7419	1,189.2581
Within Pairs	31	2,385.5000	76.9516
Total	61	38,063.2419	——

The additive property, mentioned in the first section, appears very clearly here: the degree of freedom and the sum of squares in the first two rows (for between and within pairs of twins) add up to the total shown in the last row.

There remains only the test of the hypothesis H_0. For this we calculate

$$z=\frac{1}{2}\log_e\left\{\frac{1189.2581}{76.9516}\right\}$$

$$=\frac{1}{2}\log_e\{15.4546\}=1.369$$

and refer to Fisher's tables of z with degrees of freedom $f_1=n_1=30$ and $f_2=n_2=31$. The value is clearly greater than the 1% point, so, following our rule, we reject the hypothesis to be tested. This means that the mental age of an individual is not independent of the twin pair to which she belongs; or the intraclass correlation between twins is greater than zero. An unbiased estimate of the intraclass correlation, r, is obtained by solving the equation (see Fisher (1), page 231):

$$\frac{1+r}{1-r}=\frac{1189.2581}{76.9516}=15.4546$$

In this case we find $r=0.88$.

The results of a similar analysis of the chronological age of the same group of twins are shown in Table III; in this case, of course, the differences between the members of each twin pair are zero, so the within pairs sum of squares and mean square are zero.

TABLE III

ANALYSIS OF VARIANCE OF CHRONOLOGICAL AGES OF IDENTICAL GIRL TWINS

Variance	d.f.	Sum of Squares	Mean Square
Between Pairs	30	27,871.4839	929.0495
Within Pairs	31	0	0
Total	61	27,871.4839	——

A comparison of Tables II and III shows very clearly how, in the analysis of variance, the total variation is broken up into parts which may be ascribed to different factors.

Example 2. *Reliability of Tests.*

The following examples refer to the problem of determining the reliability of a test,[9] in this case an achievement test in French Reading for Grade X. As a discussion of the theory underlying the analysis may be found in my paper (**2**), only a general outline will be given here.

The factors in which we are interested are the common ability of the group tested, the ability of each individual, the trial or practice effect and the errors of measurement. Let us assume that the score X_{st} of the t-th individual on the s-th trial of the test may be represented as a sum of these factors or components. We may express this in a mathematical form by assuming that

$$X_{st}=A+B_s+C_t+z_{st} \tag{32}$$

where $s=1, 2$; $t=1, 2, \ldots, n$. A is a measure of the common ability of the group tested and is defined as the mean for all trials and individuals; B_s is a measure of the trial or practice effect; C_t is a measure of the ability of the t-th individual; z_{st} represents the errors of measurement. Finally, since A is defined as the mean for all trials and individuals, it is necessary that

$$\left.\begin{array}{l}\sum\limits_s B_s=0\\ \sum\limits_t C_t=0\end{array}\right\} \tag{33}$$

The hypotheses which we wish to test are, (1) the hypothesis H_1, that the trial or practice effect is zero, i.e.

$$H_1: B_s=0 \tag{34}$$

and (2) the hypothesis, H_2, that the individual effect is zero, i.e.

$$H_2: C_t=0 \tag{35}$$

The latter test will tell us whether or not the mental test actually measures the abilities of the individuals.

To develop the necessary tests of these hypotheses, we first write

$$\chi^2=\sum_s\sum_t(X_{st}-A-B_s-C_t)^2 \tag{36}$$

and minimize with regard to all the quantities A, B_s, C_t, to obtain the

[9]The word "test" is used below in two distinctly different senses; it may refer to the *achievement test* or to the *test of a hypothesis*. There should be no confusion, however, as the meaning is in every case made clear by the context.

absolute minimum, χ_a^2. Assuming that H_1 and H_2 are true, separately, we minimize with regard to the remaining quantities to obtain the relative minimum values of χ^2, say χ_{r1}^2 and χ_{r2}^2, required in the test of H_1 and H_2, respectively. We find

$$\chi_a^2 = \sum_s \sum_t (X_{st} - \bar{X}_{s.} - \bar{X}_{.t} + \bar{X}_{..})^2 \tag{37}$$

$$\left.\begin{aligned} \chi_{r1}^2 &= \chi_a^2 + \chi_1^2, \text{ say} \\ \chi_{r2}^2 &= \chi_a^2 + \chi_2^2, \text{ say} \end{aligned}\right\} \tag{38}$$

where

$$\chi_1^2 = \sum_s \sum_t (\bar{X}_{s.} - \bar{X}_{..})^2 \tag{39}$$

$$\chi_2^2 = \sum_s \sum_t (\bar{X}_{.t} - \bar{X}_{..})^2 \tag{40}$$

and

$$\left.\begin{aligned} \bar{X}_{.t} &= \tfrac{1}{2} \sum_s X_{st} \\ \bar{X}_{s.} &= \frac{1}{n} \sum_t X_{st} \\ \bar{X}_{..} &= \frac{1}{2n} \sum_s \sum_t X_{st} \end{aligned}\right\} \tag{41}$$

The additive property of the sum of squares is demonstrated in the identity

$$\sum_s \sum_t (X_{st} - \bar{X}_{..})^2 \equiv \chi_1^2 + \chi_2^2 + \chi_a^2 \tag{42}$$

which may also be used, as shown in the analysis of variance table, to check the calculations.

For purposes of calculation, equations (37), (39), (40) may be written as

$$\chi_a^2 = \sum_s \sum_t X_{st}^2 - \frac{1}{n} \sum_s \left(\sum_t X_{st}\right)^2 - \tfrac{1}{2} \sum_t \left(\sum_s X_{st}\right)^2 + \frac{\left(\sum_s \sum_t X_{st}\right)^2}{2n} \tag{43}$$

$$\chi_1^2 = \frac{1}{n} \sum_s \left(\sum_t X_{st}\right)^2 - \frac{\left(\sum_s \sum_t X_{st}\right)^2}{2n} \tag{44}$$

$$\chi_2^2 = \tfrac{1}{2} \sum_t \left(\sum_s X_{st}\right)^2 - \frac{\left(\sum_s \sum_t X_{st}\right)^2}{2n} \tag{45}$$

or the first two may be written

$$\chi_a^2 = \tfrac{1}{2} \sum_t (X_{1t} - X_{2t})^2 - \frac{1}{2n} \left\{\sum_t (X_{1t} - X_{2t})\right\}^2 \tag{46}$$

$$\chi_1^2 = \frac{1}{2n} \left\{\sum_t (X_{1t} - X_{2t})\right\}^2 \tag{47}$$

and

$$\sum_s\sum_t(X_{st}-\bar{X}_{..})^2=\sum_s\sum_t X_{st}^2-\frac{1}{2n}\{\sum_s\sum_t X_{st}\}^2 \tag{48}$$

The general method used in calculating the sums of squares can be seen if we examine the items in equations (43), (44), (45) and (48). We use sums, sums of squares and sums of squared sums, i.e.

$$\begin{gathered}\sum_s\sum_t X_{st}^2\\ \sum_t X_{st} \text{ and } \sum_s(\sum_t X_{st})^2\\ \sum_s X_{st} \text{ and } \sum_t(\sum_s X_{st})^2\\ \sum_s\sum_t X_{st} \text{ and } (\sum_s\sum_t X_{st})^2\end{gathered} \tag{49}$$

or, in simple cases like the present one, sums, squares of sums, differences and squares of differences as in equations (45), (46) and (47).

To test the hypothesis H_1, we calculate

$$z_1=\tfrac{1}{2}\log_e\left\{\frac{\chi_1^2}{f_1}\Big/\frac{\chi_a^2}{f_3}\right\} \tag{50}$$

and refer to Fisher's tables of z with degrees of freedom $n_1=f_1=1$ and $n_2=f_3=n-1$.

To test the hypothesis H_2, we calculate

$$z_2=\tfrac{1}{2}\log_e\left\{\frac{\chi_2^2}{f_2}\Big/\frac{\chi_a^2}{f_3}\right\} \tag{51}$$

and refer to Fisher's tables of z with degrees of freedom

$$n_1=n_2=f_2=f_3=n-1.$$

In the following examples we have applied the method to two sets of results for the French Reading test mentioned in the first paragraph; they refer to the same test, but to results from different classes of pupils (both in Grade X). These particular examples were chosen because they illustrate nearly all the possible kinds of results which we may obtain.

Example 2 (a)

TABLE IV

SCORES RECEIVED BY PUPILS ON FORMS A AND B OF THE FRENCH READING TEST (a)

Individual, t	Score on Form A X_{1t}	Score on Form B X_{2t}	Sum $X_{1t}+X_{2t}$	Difference $X_{1t}-X_{2t}$
1	17	20	37	−3
2	18	20	38	−2
3	19	23	42	−4
4	11	7	18	4
5	19	22	41	−3
6	16	20	36	−4
7	36	38	74	−2
8	9	12	21	−3
9	12	12	24	0
10	31	34	65	−3
11	16	22	38	−6
12	10	10	20	0
13	28	30	58	−2
14	29	38	67	−9
15	35	34	69	1
16	14	18	32	−4
17	28	28	56	0
18	6	2	8	4
19	8	12	20	−4
20	18	23	41	−5
21	14	20	34	−6
22	19	20	39	−1
23	24	29	53	−5
24	7	14	21	−7
25	8	12	20	−4
26	8	13	21	−5
27	19	18	37	1
28	10	9	19	1
29	19	19	38	0
30	24	24	48	0
31	12	14	26	−2
32	25	32	57	−7
33	4	5	9	−1
34	4	9	13	−5
35	22	19	41	3
Sum	599	682	1,281	−83
Sum of Squares	12,797	16,198	57,461	529

TABLE V

Analysis of Variance of Scores on Forms A and B of the French Reading Test (a)

Variance	d.f.	Sum of Squares	Mean Square
Between Forms	1	98.4143	98.4143
Between Individuals . . .	34	5,288.2000	155.5353
Error	34	166.0857	4.8849
Total	69	5,552.7000	——

From the values in Table V we find, for the test of the hypothesis H_1,

$$z_1 = \tfrac{1}{2}\log_e\left\{\frac{98.4143}{4.8849}\right\} = 1.502$$

The degrees of freedom are $n_1 = f_1 = 1$ and $n_2 = f_3 = 34$. From Fisher's tables we find that our value of z is greater than the 1% point, so we reject the hypothesis to be tested, H_1. This means that in this case the trial or practice effect is significant.

For the test of the hypothesis H_2, we calculate

$$z_2 = \tfrac{1}{2}\log_e\left\{\frac{155.5353}{4.8849}\right\} = 1.730$$

The degrees of freedom are $n_1 = n_2 = f_2 = f_3 = 34$. Referring to Fisher's tables of z we find that our value of z is beyond the 1% point, so we reject the hypothesis H_2. We conclude that the test measures the abilities of the individuals tested, or measures them accurately enough to enable us to distinguish between the individuals.

Example 2(*b*).

The original data for this example are given in Table VI and the results of the analysis in Table VII. There are very few cases, but we may apply the method as the use of degrees of freedom makes it equally valid for large or small samples.

In the test of the hypothesis H_1, we have one of the cases discussed in the second note of the first chapter. We calculate

$$z_l = \tfrac{1}{2}\log_e\left\{\frac{12.4621}{4.6944}\right\} = 0.489$$

and refer to Fisher's tables of z with degrees of freedom $n_1 = f_3 = 17$ and

TABLE VI

SCORES RECEIVED BY PUPILS ON FORMS A AND B OF THE FRENCH READING TEST (b)

Individual i	Score on		Sum $X_{1i}+X_{2i}$	Difference $X_{1i}-X_{2i}$
	Form A X_{1i}	Form B X_{2i}		
1	23	32	55	−9
2	34	32	66	2
3	33	34	67	−1
4	31	25	56	6
5	22	29	51	−7
6	34	36	70	−2
7	35	37	72	−2
8	26	25	51	1
9	33	28	61	5
10	29	33	62	−4
11	35	26	61	9
12	26	21	47	5
13	28	28	56	0
14	34	35	69	−1
15	33	29	62	4
16	29	21	50	8
17	28	32	60	−4
18	27	24	51	3
Sum	540	527	1,067	13
Sum of Squares	16,490	15,841	64,229	433

TABLE VII

ANALYSIS OF VARIANCE OF SCORES ON FORMS A AND B OF THE FRENCH READING TEST (b)

Variance	d.f.	Sum of Squares	Mean Squares
Between Forms	1	4.6944	4.6944
Between Individuals	17	489.8056	28.8121
Error	17	211.8056	12.4621
Total	35	706.3056	——

$n_2=f_1=1$. The value of z is less than the 5% point, so we accept H_1, i.e., we conclude that the practice or trial effect is zero. We would, as a matter of fact, always accept the hypothesis H_1 in such cases, as none of the hypotheses alternative to H_1 is acceptable in this situation.

For the test of the hypothesis H_2, we calculate

$$z_2=\tfrac{1}{2}\log_e\left\{\frac{28.8121}{12.4621}\right\}=0.419$$

and refer to Fisher's tables of z with degrees of freedom $n_1=n_2=17$. Our value of z is approximately equal to the 5% point, so we should be inclined to accept the hypothesis H_2. A comparison of the results with those of Example 2(a) shows that in this case the errors of measurement are larger and the differences between individuals considerably smaller. A glance at the scores in Table VI will show that this does seem to be a selected class of students, which would explain one difference, but it is equally obvious that our test may not measure with equal accuracy in all situations.

Supplementary Notes on Theory.

(1) It should be noted that the test of the hypothesis H_1 in cases such as these is equivalent to the test of significance of the difference between the mean scores on the A and B forms, i.e. the t-test, as it is known. This test could be made separately, of course, but it is convenient to include it as one of the set of tests in the analysis of variance. The equivalence can be seen more clearly if we compare the figures in the first column ($n_1=1$) of the tables of z with the corresponding values given in the table of t; the relationship is $z=\log_e(t)$.

(2) We may obtain an estimate of the probable error of an individual score from the mean square ascribable to "error" by taking the square root and multiplying by the factor, 0.6745.

(3) The test of the hypothesis H_2 may easily be shown to be equivalent to the test of the hypothesis that the reliability coefficient, r_{AB}, is zero.

Example 3. Errors of Marking.

The data given in the following table refer to a problem we encountered in the construction of a test. The marking of one section of the test was not objective and the question arose as to whether the error introduced was large enough to justify the trouble of eliminating it by changing the form of the test or by constructing an elaborate key.

In order to determine the magnitude of the error, we chose 20 papers at random from those returned to us by a teacher and had these

marked independently by three examiners, whom we may denote by P, Q, R. The marks are shown in Table VIII, in the columns headed "Marks assigned by Examiners P, Q, R." What factors will determine the mark given by an examiner to the paper of a particular individual? The most important, obviously, should be the true value of the paper (presumably a measure of the ability of the individual), but superimposed on this there will be another factor due to the differences in the degree of severity of marking of the different examiners. Finally, there will be random differences, not ascribable to the above factors, due to the varying reactions of the examiners to different answers.

TABLE VIII

MARKS ASSIGNED BY EXAMINERS

Individual t	Marks assigned by Examiners			Sum $\sum_s X_{st}$
	P X_{1t}	Q X_{2t}	R X_{3t}	
1	18	19	19	56
2	20	20	20	60
3	30	33	31	94
4	17	18	19	54
5	24	23	22	69
6	33	32	36	101
7	21	23	22	66
8	16	18	18	52
9	33	34	35	102
10	24	25	26	75
11	12	13	14	39
12	16	15	16	47
13	23	25	24	72
14	25	27	28	80
15	30	31	32	93
16	22	25	25	72
17	24	23	28	75
18	27	29	27	83
19	27	28	31	86
20	31	30	31	92
Sum	473	491	504	1,468
Sum of Squares	11,873	12,749	13,468	114,060

If we denote by X_{st} the mark assigned by the s-th examiner to the paper of the t-th individual, we may write

$$X_{st}=A+B_s+C_t+z_{st} \tag{52}$$

where $s=1, 2, 3, ;\ t=1, 2, \ldots, n=20$. A is a measure of the common ability of the group tested; it is defined as the mean of the marks for all examiners and individuals; B_s is a measure of the differences between examiners; C_t a measure of the ability of the t-th individual, and z_{st} represents the random errors of marking. Since we have defined A as the mean of the marks for all examiners and individuals, it is necessary that

$$\left.\begin{aligned}\sum_s B_s=0\\ \sum_t C_t=0\end{aligned}\right\} \tag{53}$$

The hypotheses we wish to test are

$$\left.\begin{aligned}H_1:\ B_s=0\\ H_2:\ C_t=0\end{aligned}\right\} \tag{54}$$

If the hypothesis H_1 is true, then our problem reduces to that of eliminating, if possible, the random errors of marking. If, on the other hand, H_1 is false, we have two problems, but the solution of our original problem is simplified in either case as we shall know the kind of errors which affect the results.

If the hypothesis H_2 is true, then we know that the mark assigned to a paper bears little relation to the ability of the individual (or the true value of his paper). We knew, however, before starting this problem, that our test was a fairly accurate measuring instrument (from the results of analyses similar to that considered in Example 2), but the test of this hypothesis is included here because it is part of the general analysis and will be important in other cases.

The method used in developing the test of these hypotheses is the same as in the previous problems. We write

$$\chi^2=\sum_s\sum_t(X_{st}-A-B_s-C_t)^2 \tag{55}$$

and minimize χ^2

(1) with regard to all the quantities, to obtain the absolute minimum value of χ^2, χ_a^2;

(2) with regard to the quantities remaining when we assume H_1 is true, to obtain the first relative minimum value of χ^2, χ_{r1}^2;

(3) with regard to the quantities remaining when we assume H_2 is true, to obtain the second relative minimum value of χ^2, χ_{r2}^2.

We obtain, finally,

$$\left.\begin{aligned}A&=\frac{1}{3n}\sum_s\sum_t X_{st}=\bar{X}_{..}\\B_s&=\frac{1}{n}\sum_t X_{st}-\bar{X}_{..}=\bar{X}_{s.}-\bar{X}_{..}\\C_t&=\frac{1}{3}\sum_s X_{st}-\bar{X}_{..}=\bar{X}_{.t}-\bar{X}_{..}\end{aligned}\right\}\tag{56}$$

$$\begin{aligned}\chi_a^2&=\sum_s\sum_t(X_{st}-\bar{X}_{s.}-\bar{X}_{.t}+\bar{X}_{..})^2\\&=\sum_s\sum_t X_{st}^2-\frac{1}{n}\sum_s\Big\{\sum_t X_{st}\Big\}^2-\frac{1}{3}\sum_t\Big\{\sum_s X_{st}\Big\}^2+\frac{1}{3n}\Big\{\sum_s\sum_t X_{st}\Big\}^2\end{aligned}\tag{57}$$

$$\begin{aligned}\chi_{r1}^2&=\sum_s\sum_t(X_{st}-\bar{X}_{.t})^2\\&=\chi_a^2+\frac{1}{n}\sum_s\Big\{\sum_t X_{st}\Big\}^2-\frac{1}{3n}\Big\{\sum_s\sum_t X_{st}\Big\}^2\\&=\chi_a^2+\chi_1^2,\ \text{say.}\end{aligned}\tag{58}$$

$$\begin{aligned}\chi_{r2}^2&=\sum_s\sum_t(X_{st}-\bar{X}_{s.})^2\\&=\chi_a^2+\frac{1}{3}\sum_t\Big\{\sum_s X_{st}\Big\}^2-\frac{1}{3n}\Big\{\sum_s\sum_t X_{st}\Big\}^2\\&=\chi_a^2+\chi_2^2,\ \text{say.}\end{aligned}\tag{59}$$

To test the hypothesis H_1, we calculate

$$z_1=\tfrac{1}{2}\log_e\left\{\frac{\chi_1^2}{f_1}\Big/\frac{\chi_a^2}{f_3}\right\}\tag{60}$$

and refer to Fisher's tables of z with degrees of freedom

$$n_1=f_1=2,\ \text{and}\ n_2=f_3=2(n-1)=38.$$

To test the hypothesis H_2, we calculate

$$z_2=\tfrac{1}{2}\log_e\left\{\frac{\chi_2^2}{f_2}\Big/\frac{\chi_a^2}{f_3}\right\}\tag{61}$$

and refer to Fisher's tables of z with degrees of freedom

$$n_1=f_2=n-1=19,\ \text{and}\ n_2=f_3=2(n-1)=38.$$

The complete analysis is shown in Table IX. It will be seen that about 3% of the total variation is ascribable to differences between examiners and random errors of marking.

TABLE IX

ANALYSIS OF VARIANCE OF THE MARKS ASSIGNED BY THE DIFFERENT EXAMINERS

Variance	d.f.	Sum of Squares	Mean Square
Between Examiners . . .	2	24.2333	12.1167
Between Individuals . . .	19	2,102.9333	110.6807
Errors of Marking	38	45.7667	1.2044
Total	59	2,172.9333	———

$$z_1 = \tfrac{1}{2} \log_e \left\{ \frac{12.1167}{1.2044} \right\} = 1.154$$

$$z_2 = \tfrac{1}{2} \log_e \left\{ \frac{110.6807}{1.2044} \right\} = 2.260$$

The values of both z_1 and z_2 are significant, i.e. greater than the 1% points given in the tables of z, so we should reject both H_1 and H_2. We conclude that the mark is determined mainly by the ability of the individual (i.e. by the true value of the paper), but that part of the error in marking lies in the differing degrees of severity with which the examiners mark the papers. It was decided to construct the key in such a form as would eliminate the differences between examiners and also reduce the random errors as much as possible. These errors were not large enough, however, to make it necessary to change the form of the test itself.

***Example* 4.** *Tests of Homogeneity and the Combination of Results from Different Experiments.*

One of the problems frequently encountered in experimental work is that of determining the condition under which it is permissible to combine the results from several sub-samples. For results such as we are considering, i.e. for problems in which the analysis of variance method may be used, it is possible to do this. All the tests used in this method are based on the comparison of sums of squares or mean squares. Consequently, the conditions to be satisfied are simply that the estimates of variance ascribable to particular factors in each of the different samples, should be estimates of a variance common to all the groups. In one form, as we have seen, this assumption is basic to the method, so we are not introducing a new concept. We merely test whether or not the assumption is satisfied for groups which, because

of the nature of the experiment, may be taken as homogeneous in themselves. The problem, in a slightly different form, was first discussed by Neyman and Pearson (**9**) and investigated further by P. P. N. Nayer (**5**) and B. L. Welch (**12, 13**). Exact statistical tests are not available for cases in which there are more than 2 sub-samples, but those which are suggested by the above authors appear to be adequate.

The nature of the problem can be seen more clearly by considering an example. Wingfield (**14**) gives data relating to the measurement of the mental age of 57 pairs of fraternal twins; these may be of like- or unlike-sex, and in the like-sex pairs there are, of course, boy twins and girl twins. We have, therefore, three groups of fraternal twins and the question arises whether the results are the same for each group, i.e. whether we can speak of a common fraternal twin group and consider these three sub-groups as samples from a common population of fraternal twin pairs. The problem is not so simple, as there are several related problems, but this is its outline.

Tables X, XI, XII and XIII give the original data and the results of the analysis for the two groups of like-sex fraternal twins; girl twins in Tables X and XI, and boy twins in Tables XII and XIII. The analysis here is the same as in Example 1.

TABLE X

MENTAL AGES OF FRATERNAL LIKE-SEX TWINS—GIRLS

Twin Pairs t	Mental Age X_{1t}	X_{2t}	Sum $X_{1t}+X_{2t}$	Difference $\vert X_{1t}-X_{2t}\vert$
1	137	113	250	24
2	173	147	320	26
3	119	129	248	10
4	138	150	288	12
5	142	104	246	38
6	191	172	363	19
7	144	188	332	44
8	136	113	249	23
9	141	143	284	2
10	141	118	259	23
11	112	121	233	9
12	147	139	286	8
13	156	124	280	32
14	118	155	273	37
15	141	140	281	1
16	103	118	221	15
17	111	161	272	50
18	144	132	276	12
19	113	108	221	5
Sum	2,607	2,575	5,182	—
Sum of Squares	366,271	358,301	1,437,412	11,732

TABLE XI

ANALYSIS OF VARIANCE OF MENTAL AGES OF FRATERNAL LIKE-SEX TWINS—GIRLS

Variance	d.f.	Sum of Squares	Mean Square
Between Pairs	18	12,044.8421	669.1579
Within Pairs	19	5,866.0000	308.7368
Total	37	17,910.8421	—

TABLE XII

Mental Ages of Fraternal Like-Sex Twins—Boys

Twin Pairs t	Mental Age X_{1t}	Mental Age X_{2t}	Sum $X_{1t}+X_{2t}$	Difference $\lvert X_{1t}-X_{2t} \rvert$
1	156	176	332	20
2	101	119	220	18
3	108	121	229	13
4	134	110	244	24
5	165	174	339	9
6	158	167	325	9
7	159	137	296	22
8	130	150	280	20
9	144	129	273	15
10	173	147	320	26
11	101	97	198	4
12	113	123	236	10
Sum	1,642	1,650	3,292	—
Sum of Squares	232,162	234,100	928,992	3,532

TABLE XIII

Analysis of Variance of Mental Ages of Fraternal Like-Sex Twins—Boys

Variance	d.f.	Sum of Squares	Mean Square
Between Pairs	11	12,943.3333	1,176.6667
Within Pairs	12	1,766.0000	147.1667
Total	23	14,709.3333	——

Tables XIV, XV, XVI give the original data and the results of the analysis for the group of unlike-sex fraternal twins. There is an additional problem here, namely that of determining whether or not there is a significant difference between the mental ages of the boys and girls. The first analysis, therefore, is similar to that considered in Example 2; the results are given in Table XV.

TABLE XIV

MENTAL AGES OF UNLIKE-SEX FRATERNAL TWINS

Pairs t	Mental Ages		Sum $X_{1t}+X_{2t}$	Difference $X_{1t}-X_{2t}$
	Boys X_{1t}	Girls X_{2t}		
1	97	110	207	−13
2	129	103	232	26
3	131	139	270	− 8
4	151	132	283	19
5	180	140	320	40
6	124	118	242	6
7	133	142	275	− 9
8	158	160	318	− 2
9	128	145	273	−17
10	140	145	285	− 5
11	126	117	243	9
12	130	136	266	− 6
13	139	205	344	−66
14	130	111	241	19
15	113	138	251	−25
16	140	162	302	−22
17	177	146	323	31
18	141	144	285	− 3
19	134	135	269	− 1
20	120	127	247	− 7
21	116	116	232	0
22	138	177	315	−39
23	122	125	247	− 3
24	122	158	280	−36
25	149	164	313	−15
26	161	158	319	3
Sum	3,529	3,653	7,182	−124
Sum of Squares	487,747	526,251	2,014,668	13,328

TABLE XV

ANALYSIS OF VARIANCE OF MENTAL AGES OF UNLIKE-SEX FRATERNAL TWINS

Variance	d.f.	Sum of Squares	Mean Square
Between Sexes	1	295.6923	295.6923
Between Pairs	25	15,389.3077	615.5723
Error (Within Pairs) . . .	25	6,368.3077	254.7323
Total	51	22,053.3077	———

We find that this difference is not significant,

$$z=\tfrac{1}{2}\log_e\left\{\frac{295.6923}{254.7323}\right\}=0.075$$

$$n_1=1,\quad n_2=25$$

so we may present the results in a form similar to that for the first two groups, as shown in Table XVI.

TABLE XVI

ANALYSIS OF VARIANCE OF MENTAL AGES OF UNLIKE-SEX FRATERNAL TWINS (FINAL ANALYSIS)

Variance	d.f.	Sum of Squares	Mean Square
Between Pairs	25	15,389.3077	615.5723
Within Pairs	26	6,664.0000	256.3077
Total	51	22,053.3077	———

The analyses are based on the assumption that we may write

$$X_{sti}=A_s+C_{st}+z_{sti} \tag{62}$$

where $s=1, 2, 3$; $i=1, 2$; $t=1, 2, \ldots, n_s$. The additional subscript s is introduced to denote the particular group of fraternal twins considered. A_s is a measure of the common mental age of the s-th group, and is defined as the arithmetic mean of the mental ages for all individuals within the s-th group; C_{st} is a measure of the mental age of the t-th twin pair within the s-th group; z_{sti} is a measure of the differences between the mental ages of each twin pair within the s-th group. We must, of course, have, as before,

$$\sum_t C_{st}=0 \text{ for each } s. \tag{63}$$

We assume, in the first instance, that z_{sti} is a random term normally

distributed about zero with standard deviation σ_s, say. If the groups are samples from a common fraternal twin population, then σ_s will be the same for all s; i.e. z_{sti} will be distributed with standard deviation σ, say, constant for all groups. The first hypothesis we wish to test, therefore, may be written

$$H_1: \sigma_s = \sigma \tag{64}$$

Welch (**13**) has suggested a test of this hypothesis similar to that originally suggested by Neyman and Pearson. We calculate

$$L_1 = \pi_s \left(\frac{N}{n_s}\right)^{\frac{n_s}{N}} \pi_s \left\{\frac{\theta_s^1}{\sum_s \theta_s^1}\right\}^{\frac{n_s}{N}} \tag{65}$$

where $N = \sum_s n_s$, π denotes the product, and θ_s^1 denotes the within pairs sum of squares for the s-th group; and refer to Nayer's tables[10] of the L_1 distribution with $k=3$ and degrees of freedom $\bar{f}_2 = \frac{1}{3}\sum_s f_{s2}$, where f_{s2} denotes the degrees of freedom associated with θ_s^1 in the s-th group. The rule to be followed in using these tables is to reject the hypothesis tested when the calculated value of L_1 is *less* than the corresponding 1% point given in the table.

To find the value of L_1, we first calculate the value of log L_1 where

$$\log L_1 = \log N - \frac{1}{N}\sum_s n_s \log n_s + \frac{1}{N}\sum_s n_s \log \theta_s^1 - \log (\sum_s \theta_s^1) \tag{66}$$

and then find L_1 from a table of antilogarithms. For our case, we have:

TABLE XVII

CALCULATION OF LOG L_1 FOR THE TEST OF THE HYPOTHESIS H_1: $\sigma_s = \sigma$

f_{s2}	n_s	$\log n_s$	$n_s \log n_s$	θ_s^1	$\log \theta_s^1$	$n_s \log \theta_s^1$
26	52	1.7160	——	6,664	3.8237	——
19	38	1.5798	——	5,866	3.7683	——
12	24	1.3802	——	1,766	3.2470	——
57	114	$\sum_s n_s \log n_s = 182.3890$		14,296	$\sum_s n_s \log \theta_s^1 = 419.9590$	

We find $\bar{f}_2 = 19$, $L_1 = 0.967$, and from Nayer's tables we see that this is greater than the 5% point (0.898), so we accept the hypothesis H_1.

The other variance in which we are interested is ascribable to between pairs or, in the notation of equation (63), ascribable to the

[10](5), page 51, Tables IV and V. (See Appendix C.)

factor C_{st}. If the C_{st} are not all equal to zero, we may assume that C_{st} is normally distributed about zero with standard deviation σ_{cs}, say. If the groups are all samples from a common population of fraternal twins, then σ_{cs} will be the same for all s, i.e. $\sigma_{cs}=\sigma_c$ for all s. The hypothesis we wish to test, therefore, may be written in the form

$$H_2 \colon \sigma_{cs}=\sigma_c \tag{67}$$

We cannot test H_2 directly, unfortunately, but we can test it indirectly if we may assume H_1 is true. We consider another hypothesis H_3,

$$H_3 \colon \begin{Bmatrix} \sigma_{cs}=\sigma_c \\ \sigma_s \ =\sigma \end{Bmatrix} \tag{68}$$

and if both H_3 and H_1 are true, then H_2 must also be true.

The hypothesis H_3 is equivalent, in this case, to the one originally considered by Neyman and Pearson. If H_3 is true, than X_{sti} is normally distributed about a mean A_s with constant standard deviation; to test H_3 we calculate, following the method of Welch (13),

$$L_1 = \pi_s \left(\frac{N}{n_s}\right)^{\frac{n_s}{N}} \pi_s \left\{\frac{\theta_s}{\sum\limits_s \theta_s}\right\}^{\frac{n_s}{N}} \tag{69}$$

where θ_s denotes the total sum of squares for the s-th group; and refer to Nayer's tables of L_1 with $k=3$ and degrees of freedom $\bar{f}=\frac{1}{3}\sum\limits_s f_s$; where f_s denotes the degrees of freedom associated with θ_s in the s-th group.

Proceeding as before, we first calculate log L_1 as shown in the following table:

TABLE XVIII

CALCULATION OF LOG L_1 FOR THE TEST OF THE HYPOTHESIS H_3: $\begin{Bmatrix} \sigma_{cs}=\sigma_c \\ \sigma_s=\sigma \end{Bmatrix}$

f_s	n_s	$\log n_s$	$n_s \log n_s$	θ_s	$\log \theta_s$	$n_s \log \theta_s$
51				22,053.3077	4.3435	——
37	(See Table XVII)			17,910.8421	4.2531	——
23				14,709.3333	4.1676	——
111				54,673.4831	$\sum\limits_s n_s \log \theta_s = 487.5013$	

We find $\bar{f}=37$, $L_1=0.990$, and from Nayer's tables we see that L_1 is greater than the 5% point, so we accept H_3—and therefore H_2. We may assume that the three groups are samples from the same common population of fraternal twins, and combine the results. One further

test in the analysis and the final combination of the results will be discussed in the first example of Part 2.

It was pointed out that in the case of two samples there is an exact test of these hypotheses. In such cases, L_1 is a function of z, so we may use Fisher's z test and the tables of z in making the tests; otherwise the procedure is the same. The tests in this form are applied in the first part of the next example.

***Example* 5.** *Differences in the Mental Ages of Identical Twins.*

In this example we return to the problem raised in the first: whether the differences between twins may be due solely to the errors of measuring by means of mental tests. The following data are taken from Wingfield (**14**), and refer to the results from two mental tests given to 31 pairs of identical girl twins. From these we can find the answer to our question; the differences between the measurements of the same individuals will give us an estimate of the errors of measuring which we can use as a basis for our tests.

Before proceeding with this analysis, however, it is necessary to analyze the data separately for each mental test in order that we may make certain it is permissible to combine the results. We shall use the results deduced in the last example and, since there are only two samples, apply the exact tests of significance. The original data and the results of the analyses are shown in the following tables; the method used in the analysis is similar to that discussed in Example 1.

TABLE XIX

Mental Ages of Identical Girl Twins

1st Mental Test

Twin Pairs	Mental Age X	Mental Age Y	Sum $X+Y$	Diff. $\lvert X-Y \rvert$
1	156	163	319	7
2	120	113	233	7
3	108	115	223	7
4	112	120	232	8
5	134	120	254	14
6	131	136	267	5
7	131	138	269	7
8	124	126	250	2
9	216	158	374	58
10	135	136	271	1
11	136	151	287	15
12	121	113	234	8
13	102	112	214	10
14	136	133	269	3
15	158	197	355	39
16	114	121	235	7
17	125	127	252	2
18	138	144	282	6
19	142	158	300	16
20	120	115	235	5
21	176	175	351	1
22	145	143	288	2
23	118	123	241	5
24	225	189	414	36
25	150	163	313	13
26	141	124	265	17
27	164	154	318	10
28	145	179	324	34
29	126	121	247	5
30	131	125	256	6
31	134	156	290	22
Sum	4,314	4,348	8,662	—
Sum of Squares	622,554	626,774	2,488,932	9,724

2nd Mental Test

Twin Pairs	Mental Age X	Mental Age Y	Sum $X+Y$	Diff. $\lvert X-Y \rvert$
1	192	192	384	0
2	102	121	223	19
3	130	95	225	35
4	126	128	254	2
5	139	139	278	0
6	133	133	266	0
7	142	145	287	3
8	122	125	247	3
9	159	160	319	1
10	137	143	280	6
11	138	159	297	21
12	126	123	249	3
13	96	96	192	0
14	144	140	284	4
15	174	162	336	12
16	110	120	230	10
17	134	107	241	27
18	134	152	286	18
19	145	140	285	5
20	124	126	250	2
21	198	192	390	6
22	160	190	350	30
23	134	147	281	13
24	210	192	402	18
25	161	160	321	1
26	148	118	266	30
27	165	162	327	3
28	188	192	380	4
29	108	132	240	24
30	153	149	302	4
31	151	151	302	0
Sum	4,483	4,491	8,974	—
Sum of Squares	670,841	673,017	2,681,332	6,384

TABLE XX

ANALYSIS OF VARIANCE OF MENTAL AGES OF IDENTICAL GIRL TWINS

Variance	d.f.	Sum of Squares		Mean Square	
		Test 1	Test 2	Test 1	Test 2
Between Pairs....	30	34,300.7741	41,751.8710	1,143.3591	1,391.7290
Within Pairs......	31	4,862.0000	3,192.0000	156.8387	102.9677
Total.........	61	39,162.7741	44,943.8710	642.0127	736.7848

The within pairs and between pairs mean squares for the two tests are not significantly different, as we may show by calculating

(1) *for within pairs*

$$z=\tfrac{1}{2}\log_e\left\{\frac{156.8387}{102.9677}\right\}=0.210$$

$$n_1=n_2=31$$

(2) *for total* (see test of hypothesis H_3 in Example 4)

$$z=\tfrac{1}{2}\log_e\left\{\frac{736.7848}{642.0127}\right\}=0.070$$

$$n_1=n_2=61$$

and referring to Fisher's tables of z with the appropriate degrees of freedom. Since these are not significant, we may combine the results.

The factors present in our final analysis will be as follows:

(1) a measure of the common ability of the group tested; this we may denote by A and define as the mean for all tests and individuals;

(2) a measure of the ability of each twin pair which we may denote by C_t;

(3) a measure of the differences between the members of each pair of twins which we may denote by B_{st};

(4) a measure of the difference between the tests, denoted by D_u;

(5) the errors of measurement of mental age, denoted by z_{stu}.

If we denote by X_{stu} the score (i.e. mental age) made by the s-th individual of the t-th pair on the u-th test, we may write

$$X_{stu}=A+B_{st}+C_t+D_u+z_{stu} \tag{70}$$

where $t=1, 2, \ldots, n$; $s=1, 2$; $u=1, 2$; $N=4n$.

Since we have defined A as the mean for all individuals and all tests, it is necessary that

$$\left.\begin{aligned} \sum_u D_u &= 0 \\ \sum_t C_t &= 0 \\ \sum_s B_{st} &= 0 \end{aligned}\right\} \tag{71}$$

The hypotheses we wish to test are

$$H_1: D_u = 0 \tag{72}$$

i.e. the hypothesis that there is no difference between the tests;

and
$$H_2: C_t = 0 \tag{73}$$

i.e. the hypothesis that the score is independent of the particular twin pair to which the individual belongs;

and
$$H_3: B_{st} = 0 \tag{74}$$

i.e. the hypothesis that the difference between the members of each twin pair is zero.

Denote by

$$\chi^2 = \sum_s \sum_t \sum_u \{X_{stu} - A - B_{st} - C_t - D_u\}^2 \tag{75}$$

and minimize χ^2

(1) with regard to all the quantities, to obtain the absolute minimum value of χ^2, χ_a^2;

(2) with regard to the quantities remaining if H_1 is true, to obtain the first relative minimum value of χ^2, χ_{r1}^2;

(3) with regard to the quantities remaining if H_2 is true, to obtain the second relative minimum value of χ^2, χ_{r2}^2;

(4) with regard to the quantities remaining if H_3 is true, to obtain the third relative minimum value of χ^2, χ_{r3}^2.

We find

$$\left.\begin{aligned} A &= \frac{1}{N} \sum_s \sum_t \sum_u X_{stu} = \overline{X}_{\cdots} \\ D_u &= \frac{1}{2n} \sum_s \sum_t X_{stu} - \overline{X}_{\cdots} = \overline{X}_{\cdot\cdot u} - \overline{X}_{\cdots} \\ C_t &= \tfrac{1}{4} \sum_s \sum_u X_{stu} - \overline{X}_{\cdots} = \overline{X}_{\cdot t \cdot} - \overline{X}_{\cdots} \\ B_{st} &= \tfrac{1}{2} \sum_s X_{stu} - \tfrac{1}{4} \sum_s \sum_u X_{stu} = \overline{X}_{st\cdot} - \overline{X}_{\cdot t \cdot} \end{aligned}\right\} \tag{76}$$

$$\chi_a^2 = \sum_s \sum_t \sum_u (X_{stu} - \overline{X}_{\cdot\cdot u} - \overline{X}_{st\cdot} + \overline{X}_{\cdots})^2$$

$$=\tfrac{1}{2}\sum_s\sum_t\{X_{st1}-X_{st2}\}^2-\frac{1}{N}\{\sum_s\sum_t(X_{st1}-X_{st2})\}^2 \tag{77}$$

$$\chi_{r1}^2=\sum_s\sum_t\sum_u\{X_{stu}-\bar{X}_{st\cdot}\}^2$$

$$=\chi_a^2+\frac{1}{N}\{\sum_s\sum_t(X_{st1}-X_{st2})\}^2$$

$$=\chi_a^2+\chi_1^2,\text{ say,} \tag{78}$$

$$\chi_{r2}^2=\sum_s\sum_t\sum_u(X_{stu}-\bar{X}_{st\cdot}+\bar{X}_{\cdot t\cdot}-\bar{X}_{\cdot\cdot u})^2$$

$$=\chi_a^2+\tfrac{1}{4}\sum_t\{\sum_s\sum_u X_{stu}\}^2-\frac{1}{N}\{\sum_s\sum_t\sum_u X_{stu}\}^2$$

$$=\chi_a^2+\chi_2^2,\text{ say,} \tag{79}$$

$$\chi_{r3}^2=\sum_s\sum_t\sum_u(X_{stu}-\bar{X}_{\cdot t\cdot}-\bar{X}_{\cdot\cdot u}+\bar{X}_{\cdots})^2$$

$$=\chi_a^2+\tfrac{1}{4}\sum_t\{\sum_u(X_{1tu}-X_{2tu})\}^2$$

$$=\chi_a^2+\chi_3^2,\text{ say.} \tag{80}$$

To test the hypothesis H_1: $D_u=0$, we calculate

$$z_1=\tfrac{1}{2}\log_e\left\{\frac{\chi_1^2}{1}\Big/\frac{\chi_a^2}{2n-1}\right\} \tag{81}$$

and refer to Fisher's tables of z with degrees of freedom

$$n_1=1,\ n_2=2n-1.$$

To test the hypothesis H_2: $C_t=0$, we calculate

$$z_2=\tfrac{1}{2}\log_e\left\{\frac{\chi_2^2}{n-1}\Big/\frac{\chi_a^2}{2n-1}\right\} \tag{82}$$

and refer to Fisher's tables of z with degrees of freedom

$$n_1=n-1,\ n_2=2n-1.$$

To test the hypothesis H_3: $B_{st}=0$, we calculate

$$z_3=\tfrac{1}{2}\log_e\left\{\frac{\chi_3^2}{n}\Big/\frac{\chi_a^2}{2n-1}\right\} \tag{83}$$

and refer to Fisher's tables of z with degrees of freedom

$$n_1=n,\ n_2=2n-1.$$

The original data and the analysis for the group of identical girl twins are shown in the tables XXI and XXII ; Table XXI shows also the method of calculating the sums of squares required in the analysis and for checking the results.

TABLE XXI

Mental Ages of Identical Girl Twins (Combined Results)

Twin Pair t	Ind.	Mental Ages: Test 1 X_{st1}	Mental Ages: Test 2 X_{st2}	Diff. $X_{st1} - X_{st2}$	Sum $X_{st1} + X_{st2}$	Diff. of Sums $\lvert \sum_u (X_{1tu} - X_{2tu}) \rvert$	Sum of Sums $\sum_s \sum_u (X_{stu})$
1	1	163	192	−29	355	7	703
	2	156	192	−36	348		
2	3	113	121	−8	234	12	456
	4	120	102	18	222		
3	5	108	95	13	203	42	448
	6	115	130	−15	245		
4	7	112	126	−14	238	10	486
	8	120	128	−8	248		
5	9	134	139	−5	273	14	532
	10	120	139	−19	259		
6	11	136	133	3	269	5	533
	12	131	133	−2	264		
7	13	131	142	−11	273	10	556
	14	138	145	−7	283		
8	15	124	122	2	246	5	497
	16	126	125	1	251		
9	17	216	159	57	375	57	693
	18	158	160	−2	318		
10	19	135	137	−2	272	7	551
	20	136	143	−7	279		
11	21	136	138	−2	274	36	584
	22	151	159	−8	310		
12	23	121	126	−5	247	11	483
	24	113	123	−10	236		
13	25	102	96	6	198	10	406
	26	112	96	16	208		
14	27	136	144	−8	280	7	553
	28	133	140	−7	273		
15	29	158	174	−16	332	27	691
	30	197	162	35	359		
16	31	114	110	4	224	17	465
	32	121	120	1	241		
17	33	125	134	−9	259	25	493
	34	127	107	20	234		

TABLE XXI—*continued*

Twin Pair t	Ind.	Mental Ages Test 1 X_{st1}	Test 2 X_{st2}	Diff. $X_{st1} - X_{st2}$	Sum $X_{st1} + X_{st2}$	Diff. of Sums $\lvert\sum_u (X_{1tu} - X_{2tu})\rvert$	Sum of Sums $\sum_s \sum_u (X_{stu})$
18	35	138	134	4	272	24	568
	36	144	152	−8	296		
19	37	142	145	−3	287	11	585
	38	158	140	18	298		
20	39	120	124	−4	244	3	485
	40	115	126	−11	241		
21	41	176	198	−22	374	7	741
	42	175	192	−17	367		
22	43	145	160	−15	305	28	638
	44	143	190	−47	333		
23	45	118	134	−16	252	18	522
	46	123	147	−24	270		
24	47	225	210	15	435	54	816
	48	189	192	−3	381		
25	49	150	161	−11	311	12	634
	50	163	160	3	323		
26	51	141	148	−7	289	47	531
	52	124	118	6	242		
27	53	164	165	−1	329	13	645
	54	154	162	−8	316		
28	55	145	188	−43	333	38	704
	56	179	192	−13	371		
29	57	126	108	18	234	19	487
	58	121	132	−11	253		
30	59	131	153	−22	284	10	558
	60	125	149	−24	274		
31	61	134	151	−17	285	22	592
	62	156	151	5	307		
Sum		8,662	8,974	−312	17,636	——	17,636
Sum of Squares		1,249,328	1,343,858	18,576	5,167,796	18,600	10,316,992

TABLE XXII

Analysis of Variance of Mental Ages of Identical Girl Twins (Combined Results)

Variance	d.f.	Sum of Squares	Mean Square
Between Tests	1	785.0323	785.0323
Between Pairs of Twins . .	30	70,953.6774	2,365.1226
Within Pairs of Twins . . .	31	4,650.0000	150.0000
Error	61	8,502.9677	139.3929
Total	123	84,891.6774	———

$$z_1=\tfrac{1}{2}\log_e\left\{\frac{785.0323}{139.3929}\right\}$$

$$=0.864$$

$$f_1=n_1=1;\ f_4=n_2=61$$

$$z_2=\tfrac{1}{2}\log_e\left\{\frac{2365.1226}{139.3929}\right\}$$

$$=1.417$$

$$f_2=n_1=30;\ f_4=n_2=61$$

$$z_3=\tfrac{1}{2}\log_e\left\{\frac{150.0000}{139.3929}\right\}$$

$$=0.037$$

$$f_3=n_1=31\ ;\ f_4=n_2=61$$

Interpretation of Results.

As we see from the test of the hypothesis H_1 the difference between the two tests is probably significant; the second test gives, on the average, higher mental ages.

From the test of the hypothesis H_2, we conclude that the differences between the mental ages of the pairs of twins are significant.

From the test of the hypothesis H_3: $B_{st}=0$, we find that the differences between members of the same twin pair are not significant. These differences are of the same order of magnitude as the differences between the estimates of the mental age of the same individual from the two tests.

Comparison with Analysis of Mental Ages of Fraternal Twins (Girls).

It is interesting to compare these results with the results of a similar analysis for fraternal twins. Since the above results refer to identical girl twins, we have chosen the group of girl (i.e. like-sex) fraternal twins for this study. The original data and the results of the analysis are given in the following tables; the method used in analyzing the results and testing the hypotheses is, of course, the same as that used above.

TABLE XXIII

MENTAL AGES OF LIKE-SEX (GIRL) FRATERNAL TWINS
(COMBINED RESULTS)

Twin Pair t	Ind.	Mental Ages Test 1 X_{st1}	Mental Ages Test 2 X_{st2}	Diff. $X_{st1} - X_{st2}$	Sum $X_{st1} + X_{st2}$	Diff. of Sums $\lvert\sum_u(X_{1tu} - X_{2tu})\rvert$	Sum of Sums $\sum_s\sum_u(X_{stu})$
1	1	134	140	−6	274	49	499
	2	113	112	1	225		
2	3	159	186	−27	345	51	639
	4	153	141	12	294		
3	5	115	123	−8	238	20	496
	6	134	124	10	258		
4	7	138	137	1	275	25	575
	8	141	159	−18	300		
5	9	128	156	−28	284	76	492
	10	118	90	28	208		
6	11	189	192	−3	381	38	724
	12	177	166	11	343		
7	13	140	147	−7	287	88	662
	14	203	172	31	375		
8	15	135	137	−2	272	46	498
	16	115	111	4	226		
9	17	135	146	−11	281	4	566
	18	134	151	−17	285		
10	19	134	148	−14	282	46	518
	20	110	126	−16	236		
11	21	112	111	1	223	19	465
	22	121	121	0	242		
12	23	139	154	−15	293	15	571
	24	148	130	18	278		
13	25	146	165	−19	311	63	559
	26	116	132	−16	248		
14	27	120	116	4	236	74	546
	28	136	174	−38	310		
15	29	135	146	−11	281	2	560
	30	136	143	−7	279		
16	31	109	96	13	205	31	441
	32	121	115	6	236		
17	33	115	106	9	221	101	543
	34	167	155	12	322		
18	35	141	146	−5	287	23	551
	36	128	136	−8	264		
19	37	120	105	15	225	9	441
	38	115	101	14	216		
Sum		5,130	5,216	−86	10,346	—	10,346
Sum of Squares		709,700	738,676	8,726	2,888,026	47,046	5,729,006

TABLE XXIV

ANALYSIS OF VARIANCE OF MENTAL AGES OF LIKE-SEX (GIRLS) FRATERNAL TWINS (COMBINED RESULTS)

Variance	d.f.	Sum of Squares	Mean Square
Between Tests	1	97.3158	97.3158
Between Pairs of Twins . .	18	23,834.1842	1,324.1213
Within Pairs of Twins . . .	19	11,761.5000	619.0263
Error	37	4,265.6842	115.2888
Total	75	39,958.6842	

$$z_1 = \tfrac{1}{2} \log_e \left\{\frac{115.2888}{97.3158}\right\}$$

$$= 0.085$$

$$n_1 = f_4 = 37;\ n_2 = f_1 = 1$$

$$z_2 = \tfrac{1}{2} \log_e \left\{\frac{1324.1213}{115.2888}\right\}$$

$$= 1.221$$

$$n_1 = f_2 = 18;\ n_2 = f_4 = 37$$

$$z_3 = \tfrac{1}{2} \log_e \left\{\frac{619.0263}{115.2888}\right\}$$

$$= 0.840$$

$$n_1 = f_3 = 19;\ n_2 = f_4 = 37$$

Interpretation of Results.

From the test of the hypothesis H_1, we find that the difference between the two tests is not significant; i.e. they give, on the average, the same mental ages. From the test of the hypothesis H_2, we find again that the differences between the mental ages of the pairs of twins are significant. From the test of the hypothesis H_3, we find that the differences between members of the same twin pair are significant; i.e. they are greater than the errors of measurement of mental age. It is interesting to note, also, that the errors of measurement of mental age are of the same order of magnitude in the two cases.

PART 2

Unequal Numbers of Observations in the Classes[11]

In the first part of this chapter, we considered only examples in which there were equal numbers of observations in the different classes. Since the method may be used in other cases, we shall consider in this part the procedure to be followed when the numbers of observations in the classes are unequal. The theory underlying the tests, and their development, is the same as that considered previously, but the method to be followed in calculating the sums of squares required in the analysis is different. The difference is not great; it is mainly that the arithmetical procedure is somewhat more complicated.

***Example* 6.** *Resemblance of Fraternal Twins in Intelligence.*

In Example 4, we discussed tests of homogeneity and the combination of results from different experiments, and applied the tests to the results for three groups of fraternal twins: 26 unlike-sex pairs, 19 like-sex girl pairs and 12 like-sex boy pairs. It was found that these groups could be considered as samples from a common population of fraternal twins and that the results, therefore, could be combined. As the groups were of unequal size, the problem of combining the results was left for consideration in this section.

If we denote by X_{sti} the mental age of the i-th individual of the t-th twin pair in the s-th group, we may write

$$X_{sti} = A + B_s + C_{st} + z_{sti} \tag{84}$$

where $s = 1, 2, 3$; $i = 1, 2$; $t = 1, 2, \ldots n_s$; n_s denotes the number of twin pairs in the s-th group. A is a measure of the ability of all fraternal twins considered, and is defined as the mean of the mental ages for all groups and individuals; B_s is a measure of the ability of the s-th group; C_{st} is a measure of the ability of the t-th twin pair within the s-th group; z_{sti} is a measure of the difference between the members of the twin pairs and includes also the errors of measurement. Since A is defined as the mean mental age for all groups and individuals, it is necessary that

$$\left.\begin{array}{l} \sum\limits_s B_s = 0 \\ \sum\limits_t C_{st} = 0, \text{ for each } s \end{array}\right\} \tag{85}$$

[11] In some problems in education the data are classified on two criteria and there are unequal frequencies in the sub-classes. This type of problem has not been considered here; readers interested in these problems should consult:

Snedecor, George W., and Cox, Gertrude M. *Disproportionate Subclass Numbers in Tables of Multiple Classification.* Research Bulletin No. 180 Ames, Iowa: Iowa Agricultural Experiment Station, March 1935. Pp. 272.

The hypotheses we wish to test are

$$\left.\begin{aligned} H_1&: B_s=0 \\ H_2&: C_{st}=0 \end{aligned}\right\} \tag{86}$$

Define by

$$\chi^2=\sum_s\sum_t\sum_i\{X_{sti}-A-B_s-C_{st}\}^2 \tag{87}$$

and minimize χ^2

(1) with regard to all the quantities, to obtain the absolute minimum value of χ^2, χ_a^2;

(2) with regard to the quantities remaining if H_1 is true, to obtain the first relative minimum value of χ^2, χ_{r1}^2;

(3) with regard to the quantities remaining if H_2 is true, to obtain the second relative minimum value of χ^2, χ_{r2}^2.

We obtain

$$\left.\begin{aligned} A&=\frac{1}{N}\sum_s\sum_t\sum_i X_{sti}=\bar{X}_{\cdots} \\ &\qquad\text{where } N=2\sum_s n_s \\ B_s&=\frac{1}{2n_s}\sum_t\sum_i X_{sti}-\bar{X}_{\cdots}=\bar{X}_{s\cdot\cdot}-\bar{X}_{\cdots} \\ C_{st}&=\tfrac{1}{2}\sum_i X_{sti}-\bar{X}_{s\cdot\cdot}=\bar{X}_{st\cdot}-\bar{X}_{s\cdot\cdot} \end{aligned}\right\} \tag{88}$$

$$\begin{aligned} \chi_a^2&=\sum_s\sum_t\sum_i(X_{sti}-\bar{X}_{st\cdot})^2 \\ &=\sum_s[\tfrac{1}{2}\sum_t(X_{st1}-X_{st2})^2] \end{aligned} \tag{89}$$

$$\begin{aligned} \chi_{r1}^2&=\sum_s\sum_t\sum_i(X_{sti}-\bar{X}_{st\cdot}+\bar{X}_{s\cdot\cdot}-\bar{X}_{\cdots})^2 \\ &=\chi_a^2+\tfrac{1}{2}\sum_s\left[\frac{\{\sum_t\sum_i X_{sti}\}^2}{n_s}\right]-\frac{\{\sum_s\sum_t\sum_i X_{sti}\}^2}{N} \\ &=\chi_a^2+\chi_1^2,\quad\text{say.} \end{aligned} \tag{90}$$

$$\begin{aligned} \chi_{r2}^2&=\sum_s\sum_t\sum_i(X_{sti}-\bar{X}_{s\cdot\cdot})^2 \\ &=\chi_a^2+\tfrac{1}{2}\sum_s\left[\sum_t\{X_{st1}+X_{st2}\}^2-\frac{\{\sum_t\sum_i X_{sti}\}^2}{n_s}\right] \\ &=\chi_a^2+\chi_2^2,\quad\text{say.} \end{aligned} \tag{91}$$

To test the hypothesis H_1, we calculate

$$z_1=\tfrac{1}{2}\log_e\left\{\frac{\chi_1^2}{f_1}\Big/\frac{\chi_a^2}{f_3}\right\} \tag{92}$$

and refer to Fisher's tables of z with degrees of freedom

$$n_1=f_1=2,\ n_2=f_3=\sum_s n_s=57.$$

To test the hypotheses H_2, we calculate

$$z_2=\tfrac{1}{2}\log_e\left\{\frac{\chi_2^2}{f_2}\Big/\frac{\chi_a^2}{f_3}\right\} \tag{93}$$

and refer to Fisher's tables of z with degrees of freedom

$$n_1=f_2=\sum_s(n_s-1)=54, \text{ and } n_2=f_3=\sum_s n_s=57.$$

We see from equations (89) and (91) that we can obtain the values of χ_a^2 and χ_2^2 from the analysis given in Example 4; the sum of the three "within pairs" sums of squares will give us the value of χ_a^2, and the sum of the three "between pairs" sums of squares will give us the value of χ_2^2. The value of χ_1^2, however, must be calculated from the totals given in Tables X, XII, XIV. In this calculation, we have to divide each square, $(\sum_t\sum_i X_{sti})^2$, by n_s before summing for s: this is the difference in procedure mentioned above. We find

$$\chi_1^2=\tfrac{1}{2}\left[\frac{(5182)^2}{19}+\frac{(3292)^2}{12}+\frac{(7182)^2}{26}\right]$$

$$-\frac{(15656)^2}{114}=67.8502$$

The complete analysis in the usual form is given in the following table.

TABLE XXV

Analysis of Variance of Mental Ages of Fraternal Twins (Combined Results)

Variance	d.f.	Sum of Squares	Mean Square
Between groups	2	67.8502	33.9251
Between pairs	54	40,377.4831	747.7312
Within pairs	57	14,296 0000	250.8070
Total	113	54,741.3333	----

To test H_1, we calculate

$$z_1=\tfrac{1}{2}\log_e\left\{\frac{250.8070}{33.9251}\right\}=1.00$$

and refer to Fisher's tables of z with degrees of freedom $n_1=57$ and $n_2=2$ (see Note 2, chapter 1, page 16). We find z_1 is less than the 5% point so we accept H_1 and conclude that the differences between the average ability of the groups are not significant. It is not necessary to calculate z_1 in this case, however, as it is clear from the values of the mean squares that we should accept H_1 (see also Example 2(*b*)).

To test the hypothesis H_2, we calculate

$$z_2=\tfrac{1}{2}\log_e\left\{\frac{747.7312}{250.8070}\right\}=0.546$$

and refer to Fisher's tables of z with degrees of freedom $n_1=54$ and $n_2=57$. Since $z_2=0.546$ is greater than the corresponding 1% point, we reject the hypothesis H_2. We conclude that the differences between twin pairs are significant or, as in Example 1, that the intraclass correlation is greater than zero.

Example 7. Comparison of the Ability of Pupils in Different Classes in the Same Grade.

Since there are generally more pupils enrolled in each grade in a High School than can be accommodated in a single classroom, it is necessary to divide the pupils into classes. If we know that the classes are equal in ability and variability, then we may make a direct comparison of their progress, but if this is not known we must, before making the comparison, determine the relative ability and variability of each class. The present example, therefore, is concerned with the problem of determining whether the ability and variability of different classes are equal.

The data used in the two illustrative cases considered here refer to the scores made on an intelligence test which was given to the pupils in two grades, IX and X, in an Ontario High School; there were six classes in Grade IX and five in Grade X. Since 343 pupils were tested altogether, the tables giving all the original scores would be too bulky to be included here, but the relevant sums and sums of squares for each class and grade are given in the following tables.

TABLE XXVI

Sums and Sums of Squares of Scores for Each Class (Grade IX)

Class	Number of Pupils $[n_s]$	Sum of Scores $\sum_t X_{st}$	Sum of Squares of Scores $\sum_t X_{st}^2$	Sum of Squares about Means $\sum_t X_{st}^2 - \frac{(\sum_t X_{st})^2}{n_s}$
A	32	530	10,770	1,991.8750
B	34	704	17,550	2,973.0588
C	31	446	7,210	793.3548
D	34	792	21,556	3,107.0588
E	25	520	12,336	1,520.0000
F	30	510	9,518	848.0000
Total	186	3,502	78,940	11,233.3474

TABLE XXVII

Sums and Sums of Squares of Scores for Each Class (Grade X)

Class	Number of Pupils $[n_s]$	Sum of Scores $\sum_t X_{st}$	Sum of Squares of Scores $\sum_t X_{st}^2$	Sum of Squares about Means $\sum_t X_{st}^2 - \frac{(\sum_t X_{st})^2}{n_s}$
A	33	928	28,030	1,933.5152
B	28	760	22,750	2,121.4286
C	31	1,013	35,287	2,184.7742
D	34	1,335	56,637	4,218.6176
E	31	748	21,336	3,287.4839
Total	157	4,784	164,040	13,745.8195

If we denote by X_{st} the score made by the t-th pupil in the s-th class, the basic assumption in the analysis is that we may write

$$X_{st} = A + B_s + z_{st} \tag{94}$$

where $s = 1, 2, \ldots, k$; $t = 1, 2, \ldots, n_s$; n_s denotes the number of pupils in the s-th class and k denotes the number of classes. A is a measure of the ability of all the pupils in the grade, and is defined as the mean score for all individuals and classes; B_s is a measure of the ability of the s-th class; z_{st} is a measure of the differences within classes;

it is assumed to be normally distributed about zero with constant standard deviation σ. It follows from the definition of A that

$$\sum_s B_s = 0 \tag{95}$$

The hypothesis we wish to test, namely, that the classes are of equal ability, may be written

$$H_1 : B_s = 0 \tag{96}$$

If we assume that the standard deviation σ is constant, we are assuming that the variability of the scores is the same for each class. This assumption may not be satisfied in practice, and, in fact, difference in variability is one of the factors in which we are interested, so we must first test the hypothesis

$$H_0 : \sigma_s = \sigma \tag{97}$$

where σ_s denotes the standard deviation of the scores in the s-th class. If this hypothesis is accepted, we conclude that there is no difference in variability and proceed to the test of the hypothesis H_1. If, on the other hand, we reject the hypothesis H_0, we conclude that the classes differ in variability and, in addition, we find that we cannot make an exact test of the hypothesis H_1. This is one limitation of the analysis of variance method, but in this particular case it does not matter much, from the practical point of view, as we cannot directly compare the progress of classes of different variability.

The test of the hypothesis H_0 may be made by using the method discussed in Examples 4 and 5. The computation of L_1, for the six classes in Grade IX, may be carried out as shown in the following table.

TABLE XXVIII

CALCULATION OF L_1 FOR THE TEST OF THE HYPOTHESIS H_0: $\sigma_s = \sigma$ (GRADE IX)

n_s	$\log n_s$	$n_s \log n_s$	θ_s	$\log \theta_s$	$n_s \log \theta_s$
32	1.5052	——	1,991.8750	3.2993	——
34	1.5315	——	2,973.0588	3.4732	——
31	1.4914	——	793.3548	2.8995	——
34	1.5315	——	3,107.0588	3.4923	——
25	1.3979	——	1,520.0000	3.1818	——
30	1.4771	——	848.0000	2.9284	——
186	$\sum_s n_s \log n_s = 277.7997$		11,233.3474	$\sum_s n_s \log \theta_s = 599.6866$	

We obtain $L_1 = 0.890$, and referring to Nayer's tables with $k = 6$ and mean n_s of 31 we find that L_1 is less than the 1% point. We reject the hypothesis H_0 and conclude that the classes are of different variability; we cannot make a further analysis, therefore, to test the hypothesis H_1 in this case.

Turning to the results for Grade X, the computation of L_1, to test the hypothesis H_0 for the five classes in this grade, may be carried out as shown in the following table.

TABLE XXIX

CALCULATION OF L_1 FOR THE TEST OF THE HYPOTHESIS $H_0: \sigma_s = \sigma$
(GRADE X)

n_s	$\log n_s$	$n_s \log n_s$	θ_s	$\log \theta_s$	$n_s \log \theta_s$
33	1.5185	——	1,933.5152	3.2863	——
28	1.4472	——	2,121.4286	3.3266	——
31	1.4914	——	2,184.7742	3.3394	——
34	1.5315	——	4,218.6176	3.6252	——
31	1.4914	——	3,287.4839	3.5169	——
157	$\sum_s n_s \log n_s = 235.1661$		13,745.8195	$\sum_s n_s \log \theta_s = 537.3952$	

We obtain $L_1 = 0.961$, and referring to Nayer's tables with $k = 5$ and mean n_s of 31.4 we find that L_1 is greater than the 5% point. We accept H_0, and conclude that the classes are of equal variability.

Since our assumption of constant standard deviation, $\sigma_s = \sigma$, is satisfied, we may proceed to the test of the hypothesis H_1. The development of the test is similar to that considered in Example 1 and will not be repeated here; the only difference, in fact, is that in this case we have unequal numbers of observations in the different classes. The complete analysis is given in Table XXX.

TABLE XXX

ANALYSIS OF VARIANCE OF SCORES IN DIFFERENT CLASSES IN GRADE X

Variance	d.f.	Sum of Squares	Mean Square
Between Classes	4	4,519.3015	1,129.8254
Within Classes	152	13,745.8195	90.4330
Total	156	18,265.1210	———

The within classes sum of squares may be obtained directly from the last row of Table XXIX, $\sum_{s}\theta_s = 13{,}745.8195$; the between means of classes, or between classes for short, sum of squares is calculated from the totals given in the third column of Table XXVII as follows (see explanation in Example 6)

$$\frac{(1335)^2}{34}+\frac{(1013)^2}{31}+\frac{(928)^2}{33}+\frac{(748)^2}{31}+\frac{(760)^2}{28}-\frac{(4784)^2}{157}=4{,}519.3015.$$

The total sum of squares (to be used in checking the calculations) is

$$164040-\frac{(4784)^2}{157}=18{,}265.1210$$

To test the hypothesis H_1, we calculate (from Table XXX)

$$z=\tfrac{1}{2}\log_e\left\{\frac{1{,}129.8254}{90.4330}\right\}=1.263$$

and refer to Fisher's tables of z with degrees of freedom $n_1=4$ and $n_2=152$. Since z is greater than the 1% point, we reject the hypothesis H_1 and conclude that the classes differ in ability.

In the first case (Grade IX), we found that the classes differed in variability, and probably in ability also but we could not make an exact test of the significance of these latter differences. In the second case (Grade X), we found that the classes were of equal variability but differed significantly in ability. A comparison of the progress of these classes, therefore, would be valid only if these factors were taken into consideration.

***Example* 8.** *Test of the Linearity of Regression.*

The statistical method of regression, linear and non-linear, is widely used in education and other fields. In every case the difficulty arises of determining whether or not the particular type of regression considered adequately represents the observed data. The analysis of variance technique gives us a very simple solution of the problem, and the method is applicable to both linear and non-linear regression. As the most important case in practice is the straight regression line, we shall discuss in this example only the application of the method to the problem of testing the linearity of regression. The method will be used in analyzing the results from two experiments; in the first the regression is linear but in the second it is not.

It should be pointed out that underlying the statistical method of

correlation, which is so widely used in education, are the assumptions that, (1) the regressions are linear in form, i.e. that the relationship between the variables can be adequately represented by a straight line, and (2) the variances of the different arrays are equal. As the validity of these assumptions should be tested for every case, the methods discussed below should be of particular value in educational statistics.

In the case considered here, the practical problem was that of determining whether the relationship between the scores of the same individuals on two mental tests was linear in form. If we denote by X and y the scores on the first and second tests, respectively, then, if the regression is linear, the regression function may be written

$$Y=a+b(X-\bar{X}) \tag{98}$$

where a and b are two parameters to be estimated from the data; b is the regression coefficient of y on X, Y is the expected value of y for each X, and $\bar{X}$ is the mean value of the distribution of X. The same methods may be used, of course, in studying the form of the regression of X on y.

For each selected value of X we shall generally have several values of y; these form what is termed an array. If the data are grouped, as in our case, the mid-points of the class-intervals may be taken as the selected values of X. If we denote by y_{st} the score (on the second test) of the t-th individual in the s-th array, we may write

$$Y_{st}=A+B_s+z_{st} \tag{99}$$

where $t=1, 2, \ldots, n_s$; $s=1, 2, \ldots, k$; k is the number of arrays and n_s is the number of individuals (i.e. observations) in the s-th array. A is a measure of the general ability of the individuals in the group tested and is defined as the mean of the scores, on the second test, for all individuals and arrays; B_s is a measure of the ability of the individuals in the s-th array, i.e. as determined by the scores on the second test; z_{st} is a measure of the errors of measurement by means of the second test, including errors of grouping, etc., which are independent of the value of X. It is assumed to be a random variable normally distributed about zero with the standard deviation σ, say, constant for all arrays.

The analysis here, therefore, is similar to that considered in Example 7; the total sum of squares may be broken up into two parts, one giving us the variance between means of arrays and the other the variance within arrays. The relevant data are given, in summary form, in the following tables. Before proceeding with our analysis,

TABLE XXXI

Case I. Sum and Sum of Squares of Scores in Each Array

Array No.	Value of X_s	n_s	Sum of Scores $\sum_t y_{st}$	Sum of Squares of Scores $\sum_t y_{st}^2$	$\frac{(\sum_t y_{st})^2}{n_s}$	$\sum_t y_{st}^2 - \frac{(\sum_t y_{st})^2}{n_s}$
1	8.5	19	249	3,533	3,263.2105	269.7895
2	10.5	20	311	5,729	4,836.0500	892.9500
3	12.5	32	613	12,429	11,742.7813	686.2187
4	14.5	35	727	16,293	15,100.8286	1,192.1714
5	16.5	34	711	16,299	14,868.2647	1,430.7353
6	18.5	40	956	24,444	22,848.4000	1,595.6000
7	20.5	52	1,314	34,748	33,203.7692	1,544.2308
8	22.5	38	1,066	31,472	29,904.1053	1,567.8947
9	24.5	29	885	27,891	27,007.7586	883.2414
10	26.5	35	1,085	34,551	33,635.0000	916.0000
11	28.5	39	1,267	42,405	41,161.2564	1,243.7436
12	30.5	29	1,062	39,540	38,891.1724	648.8276
13	32.5	22	850	33,734	32,840.9091	893.0909
14	34.5	30	1,171	46,597	45,708.0333	888.9667
15	36.5	19	809	35,385	34,446.3684	938.6316
16	38.5	17	714	30,532	29,988.0000	544.0000
17	40.5	21	979	46,371	45,640.0476	730.9524
18	42.5	19	950	47,982	47,500.0000	482.0000
19	44.5	14	707	36,315	35,703.5000	611.5000
20	46.5	16	842	44,740	44,310.2500	429.7500
21	48.5	22	1,215	67,903	67,101.1364	801.8636
22	50.5	18	1,064	63,[illegible]80	62,894.2222	585.7778
23	52.5	14	912	59,940	59,410.2857	529.7143
Total		614	20,459	802,313	782,005.3497	20,307.6503

however, we must test the validity of our assumption that the standard deviation σ is constant for all arrays.

We may test the hypothesis that σ is constant for all arrays by the method used in the previous example. For the first case, we find

TABLE XXXII

CASE 2. SUM AND SUM OF SQUARES OF SCORES IN EACH ARRAY

Array No.	Value of X_s	n_s	Sum of Scores $\sum_t y_{st}$	Sum of Squares of Scores $\sum_t y_{st}^2$	$\frac{(\sum_t y_{st})^2}{n_s}$	$\sum_t y_{st}^2 - \frac{(\sum_t y_{st})^2}{n_s}$
1	6.5	18	205	2,723	2,334.7222	388.2778
2	8.5	16	194	2,720	2,352.2500	367.7500
3	10.5	17	231	3,365	3,138.8824	226.1176
4	12.5	22	313	4,905	4,453.1364	451.8636
5	14.5	27	394	6,256	5,749.4815	506.5185
6	16.5	22	386	7,136	6,772.5455	363.4545
7	18.5	32	563	10,501	9,905.2813	595.7187
8	20.5	24	503	11,171	10,542.0417	628.9583
9	22.5	34	681	14,395	13,640.0294	754.9706
10	24.5	34	782	18,704	17,986.0000	718.0000
11	26.5	46	1,028	24,548	22,973.5652	1,574.4348
12	28.5	35	901	24,023	23,194.3143	828.6857
13	30.5	33	900	26,022	24,545 4545	1,476.5455
14	32.5	30	872	26,452	25,346 1333	1,105.8667
15	34.5	25	760	23.598	23,104.0000	494.0000
16	36 5	31	1,023	34,555	33,759.0000	796.0000
17	38.5	17	607	22,663	21,673.4706	989.5294
18	40.5	21	766	28,352	27,940.7619	411.2381
19	42 5	22	845	33,405	32,455.6818	949.3182
20	44.5	15	635	27,509	26,881.6667	627.3333
21	46.5	20	900	41,672	40,500.0000	1,172.0000
22	48.5	21	960	44,348	43,885.7143	462.2857
23	50.5	16	781	38,585	38,122.5625	462.4375
24	52.5	19	1,006	54,262	53,265.0526	996.9474
25	54.5	17	1,061	66,961	66,218.8824	742.1176
Total		614	17,297	598,831	580,740.6305	18,090.3695

$L_1 = 0.970$ which, for $k=23$ and mean array frequency $\bar{n}_s = 27$, is greater than the 5% point, so we may assume that σ is constant. The first analysis of the scores for this case is given in the following table:

TABLE XXXIII

ANALYSIS OF VARIANCE OF SCORES ON SECOND TEST (CASE I)

Variance	d.f.	Sum of Squares	Mean Square
Between Means of Arrays .	22	100,294.1429	4,558.8247
Within Arrays	591	20,307.6503	34.3615
Total	613	120,601.7932	———

The sums of squares are obtained from the totals given in Table XXXI. The within arrays sum of squares is simply the total of the last column; the between means of arrays sum of squares is calculated from the totals of columns (4) and (6), i.e.

$$782{,}005.3497-\frac{(20459)^2}{614}=100{,}294.1429$$

and the total sum of squares from the totals of columns (4) and (5), i.e.

$$802{,}313-\frac{(20459)^2}{614}=120{,}601.7932$$

The test of the hypothesis of a constant σ is similar for the second case; we find $L_1=0.926$, $k=25$ and $\bar{n}_s=24.6$, which is approximately equal to the value given for the 5% point, so again we may assume that σ is constant. The first analysis of variance of the scores for this case is given below; the sums of squares are calculated as explained above.

TABLE XXXIV

ANALYSIS OF VARIANCE OF SCORES ON SECOND TEST (CASE 2)

Variance	d.f.	Sum of Squares	Mean Square
Between Means of Arrays .	24	93,466.6745	3,894.4447
Within Arrays	589	18,090.3695	30.7137
Total	613	111,557.0440	———

The hypothesis we wish to test is that the regression of y on x is linear, i.e.

$$H_1: Y_s=a+b(X_s-\bar{X}) \tag{100}$$

where Y is the expected value of y for the s-th value of X. If this hypothesis is true, we may write $a+b(X_s-X)$ instead of $A+B_s$ in

equation (99), and the test of H_1 may be developed as in the previous examples. We write

$$\chi^2 = \sum_s \sum_t (y_{st} - A - B_s)^2 \tag{101}$$

and minimize χ^2 with regard to all the parameters to obtain the absolute minimum value of χ^2, χ_a^2.

We have

$$\chi_a^2 = \sum_s \sum_t y_{st}^2 - \sum_s \left\{ \frac{(\sum_t y_{st})^2}{n_s} \right\} \tag{102}$$

which is the within arrays sum of squares. The second step is to minimize χ^2 with regard to the parameters remaining if H_1 is true, i.e. minimize

$$\chi^2 = \sum_s \sum_t \{ y_{st} - a - b(X_s - \bar{X}) \}^2 \tag{103}$$

with regard to the parameters a and b to obtain the relative minimum value of χ^2, χ_r^2. We obtain

$$a = \frac{1}{\sum_s n_s} \sum_s \sum_t y_{st} = \bar{y}_{..} \tag{104}$$

$$b = \frac{\sum_s [(X_s - \bar{X})(\sum_t y_{st})]}{\sum_s \{ n_s (X_s - \bar{X})^2 \}} \tag{105}$$

$$\chi_r^2 = \chi_a^2 + \sum_s \left\{ \frac{(\sum_t y_{st})^2}{n_s} \right\} - \frac{(\sum_s \sum_t y_{st})^2}{\sum_s n_s} - \frac{[\sum_s \{ (X_s - X)(\sum_t y_{st}) \}]^2}{\sum_s \{ n_s (X_s - \bar{X})^2 \}} \tag{106}$$

$$= \chi_a^2 + \chi_1^2, \quad \text{say.} \tag{107}$$

χ_1^2 is seen to be equal to the sum of squares ascribable to between means of arrays minus the quantity

$$l = \frac{[\sum_s \{ (X_s - \bar{X})(\sum_t y_{st}) \}]^2}{\sum_s \{ n_s (X_s - \bar{X})^2 \}} \tag{108}$$

To test the hypothesis H_1, we calculate

$$z = \tfrac{1}{2} \log_e \left\{ \frac{\chi_1^2}{n_1} \Big/ \frac{\chi_a^2}{n_2} \right\} \tag{109}$$

and refer to Fisher's tables of z with degrees of freedom $n_1 = k - 2$ and $n_2 = \sum_s n_s - k$. It is convenient, and also customary, to present the analysis in the form shown in the following tables; l is entered as the "variance due to linear regression" and χ_1^2 as the "variance due to departures from linear regression".

TABLE XXXV

ANALYSIS OF VARIANCE OF SCORES ON SECOND TEST (CASE 1)—
COMPLETE ANALYSIS

Variance	d.f.	Sum of Squares	Mean Square
Due to Linear Regression .	1	99,203.3378	99,203.3378
Due to Departures from Linear Regression . .	21	1,090.8051	51.9431
Within Arrays	591	20,307.6503	34.3615
Total	613	120,601.7932	———

For the test of the hypothesis H_1, in Case 1, we find

$$z = \frac{1}{2} \log_e \left\{ \frac{51.9431}{34.3615} \right\} = 0.207$$

and we see from Fisher's tables, entered with degrees of freedom $n_1 = 21$ and $n_2 = 591$, that this value is less than the 5% point of the distribution of z. We accept the hypothesis H_1 and conclude that the regression is linear.

TABLE XXXVI

ANALYSIS OF VARIANCE OF SCORES ON SECOND TEST (CASE 2)—
COMPLETE ANALYSIS

Variance	d.f.	Sum of Squares	Mean Square
Due to Linear Regression	1	89,034.3367	89,034.3367
Due to Departures from Linear Regression . .	23	4,432.3378	192.7103
Within Arrays	589	18,090.3695	30.7137
Total	613	111,557.0440	———

For the test of the hypothesis H_1, in Case 2, we obtain

$$z = \frac{1}{2} \log_e \left\{ \frac{192.7103}{30.7137} \right\} = 0.918$$

We find from Fisher's tables of z, entered with degrees of freedom $n_1 = 23$ and $n_2 = 589$, that this value is greater than the 1% point. We reject the hypothesis H_1, therefore, and conclude that the regression is non-linear in form. It will be noticed, however, that the greater part of the variance is accounted for by the linear regression factor.

CHAPTER III

ANALYSIS OF COVARIANCE

It was pointed out in the first chapter that the analysis of covariance is an extension of the analysis of variance method. If we are considering two variates, for example, we may make a separate analysis of the variance of each and if the two variables are related we may make, at the same time, an analysis of the covariance. In the analysis of covariance we break up the sum of products into parts ascribable to different factors in much the same way as we break up the sum of squares in the analysis of variance. In the examples which follow, we shall consider only the case of two variates, but the method may easily be extended to cover the case of three or more.

Example 9. *Resemblance of Fraternal Twins (Unlike-Sex Pairs) in Intelligence.*

In Example 4, we discussed the analysis of the mental age scores of fraternal twins, and in Example 1 we showed that a similar analysis could be made for chronological age. In this example we shall consider the combined analysis of mental age and chronological age in the case of the fraternal twins (unlike-sex pairs). The analysis of the mental age scores for this group may be found in Example 4, Table XVI.

Denote by X_{it} the mental age of the i-th member of the t-th twin pair, and by Y_{it} the chronological age of the i-th member of the t-th twin pair. Following the method of Example 1, we write

$$X_{it}=A+C_t+z_{it} \tag{110}$$

and

$$Y_{it}=B+D_t \tag{111}$$

where $i=1, 2$; $t=1, 2, \ldots, n$; where n is the number of twin pairs. In the case of chronological age, of course, the members of the same twin pair are of the same chronological age, i.e. $Y_{1t}=Y_{2t}=Y_t$, so the measure of the differences within pairs of twins is zero. The total variation in chronological age, therefore, is ascribable to differences between pairs of twins.

The differences between the mental ages of the pairs of fraternal twins may be due partly or wholly to differences in chronological age. Our problem, therefore, is to determine what proportion of these differences may properly be ascribable to differences in chronological age, and adjust our analysis accordingly. If we may assume that the relationship between mental age and chronological age is linear in form, then, since Y_t denotes the chronological age of the t-th twin pair, we may write

$$X_{it}=a+bY_t+S_{it} \tag{112}$$

where a and b are constants to be estimated from the data; b is the regression coefficient of mental age on chronological age, and S_{it} is a measure of the differences between the mental ages of members of the same twin pair and differences between the mental ages of pairs of twins not ascribable to the factor of chronological age.

Following the usual method, we minimize

$$\chi^2=\sum_i\sum_t\{X_{it}-a-bY_t\}^2 \tag{113}$$

with regard to a and b to obtain the relative minimum value of χ^2, χ_r^2, say. We have

$$a=\frac{1}{2n}\left[\sum_i\sum_t X_{it}-b\sum_i\sum_t Y_t\right] \tag{114}$$

$$b=\frac{\sum_t\left[Y_t\left(\sum_i X_{it}\right)\right]-\frac{\left(\sum_t Y_t\right)\left(\sum_i\sum_t X_{it}\right)}{n}}{2\left[\sum_t Y_t^2-\frac{\left(\sum_t Y_t\right)^2}{n}\right]} \tag{115}$$

$$\chi_r^2=\tfrac{1}{2}\sum_t(X_{1t}-X_{2t})^2+\tfrac{1}{2}\left[\sum_t(X_{1t}+X_{2t})^2-\frac{\left(\sum_i\sum_t X_{it}\right)^2}{n}\right]$$

$$-\frac{\left[\sum_t\left\{Y_t\left(\sum_i X_{it}\right)\right\}-\frac{\left(\sum_t Y_t\right)\left(\sum_i\sum_t X_{it}\right)}{n}\right]^2}{2\left[\sum_t Y_t^2-\frac{\left(\sum_t Y_t\right)^2}{n}\right]}$$

$$=\chi_a^2+\chi_1^2,\quad \text{say.} \tag{116}$$

The portion of the variance ascribable to chronological age, therefore, is

$$l = \frac{\left[\sum_t \left\{ Y_t \left(\sum_i X_{it}\right)\right\} - \frac{\left(\sum_t Y_t\right)\left(\sum_i \sum_t X_{it}\right)}{n}\right]^2}{2\left[\sum_t Y_t^2 - \frac{\left(\sum_t Y_t\right)^2}{n}\right]} \tag{117}$$

since the other two quantities in (116) are simply the within and between pairs sums of squares for mental age. To obtain χ_1^2 we subtract l from the between pairs sums of squares.

The necessary calculations may easily be made if the data are arranged as shown in the following table.

TABLE XXXVII

MENTAL AND CHRONOLOGICAL AGES OF UNLIKE-SEX FRATERNAL TWINS

Twin Pair	Mental Age (Mos.) X_{ii}	Difference $\lvert X_{1i} - X_{2i} \rvert$	Sum $X_{1i} + X_{2i}$	Chronological Age (Mos.) Y_i
1	97 110	13	207	116
2	103 129	26	232	121
3	139 131	8	270	135
4	151 132	19	283	176
5	180 140	40	320	158
6	118 124	6	242	111
7	142 133	9	275	142
8	158 160	2	318	154
9	145 128	17	273	135
10	145 140	5	285	177
11	126 117	9	243	128
12	130 136	6	266	150
13	139 205	66	344	140
14	130 111	19	241	108

TABLE XXXVII—*Continued*

Twin Pair	Mental Age (Mos.) X_{it}	Difference $\lvert X_{1t}-X_{2t} \rvert$	Sum $X_{1t}+X_{2t}$	Chronological Age (Mos.) Y_t
15	138 113	25	251	126
16	162 140	22	302	145
17	146 177	31	323	160
18	144 141	3	285	132
19	135 134	1	269	154
20	120 127	7	247	135
21	116 116	0	232	144
22	177 138	39	315	164
23	125 122	3	247	128
24	122 158	36	280	146
25	149 164	15	313	152
26	161 158	3	319	136
Total	7,182	—	7,182	3,673
Sum of Squares	1,013,998	13,328	2,014,668	526,963

The results are generally presented in the form shown below.

TABLE XXXVIII
ANALYSIS OF VARIANCE AND COVARIANCE OF MENTAL AND CHRONOLOGICAL AGE SCORES

Variance	d.f.	Sum of Squares (Chronological Age)	Sum of Products (C.A.) (M.A.)	Sum of Squares (Mental Age)
Between Pairs of Twins	25	16,162.231	9,693.385	15,389.308
Within Pairs of Twins	26	0	0	6,664.000
Total	51	16,162.231	9,693.385	22,053.308

The quantities entered in the above table are calculated as follows:

Mental Age

$$\text{Between Pairs: } \frac{2014668}{2} - \frac{(7182)^2}{52} = 15{,}389.308$$

$$\text{Within Pairs: } \frac{13328}{2} = 6{,}664.000$$

$$\text{Total: } 1013998 - \frac{(7182)^2}{52} = 22{,}053.308$$

Chronological Age

$$\text{Between Pairs: } 2(526963) - \frac{\{(2)(3673)\}^2}{52} = 16{,}162.231$$

$$\text{Within Pairs: } \quad 0$$

$$\text{Total: } 1053926 - \frac{(7346)^2}{52} = 16{,}162.231$$

Products (C.A.)(M.A.)

$$\text{Between Pairs: } \{(116)(207) + (121)(232) + \ldots\ldots$$
$$\ldots\ldots + (136)(319)\} - \frac{(7182)(3673)}{26} = 9{,}693.385$$

$$\text{Within Pairs: } \quad 0$$

$$\text{Total: } \{(116)(97) + (116)(110) + \ldots + (136)(161)$$
$$+ (136)(158)\} - \frac{(7182)(7346)}{52} = 9{,}693.385$$

It will be seen that the between pairs sum of products gives us the quantity inside the brackets in the numerator of (117), and the denominator in the same equation may be obtained from the corresponding entry in the chronological age column. We find

$$l=\frac{(9693.385)^2}{16162.231}=5,813.660$$

The sums of squares for the analysis of mental age corrected for the effect of chronological age; or the adjusted sums of squares and mean squares as they are termed, are shown in the following table.

TABLE XXXIX

Analysis of Variance of Mental Age Scores of Fraternal Twins—Original and Adjusted Sums of Squares and Mean Squares

Variance	Original Analysis			Adjusted Analysis		
	d.f.	Sum of Squares	Mean Square	d.f.	Sum of Squares	Mean Square
Between Pairs . .	25	15,389.308	615.572	24	9,575.648	398.985
Due to effect of Chronological Age .	—	——	——	1	5,813.660	5,813.660
Within Pairs . . .	26	6,664.000	256.308	26	6,664.000	256.308
Total	51	22,053.308	——	51	22,053.308	——

The adjusted between pairs sum of squares and mean square give us, as we see from (116), a measure of the differences between twin pairs in mental age freed from the influence of chronological age. To test the hypothesis that these adjusted differences are zero, we calculate

$$z=\tfrac{1}{2}\log_e\left\{\frac{X_1^2}{n_1}\bigg/\frac{X_a^2}{n_2}\right\} \tag{118}$$

and refer to Fisher's tables of z with degrees of freedom $n_1=n-2$ and $n_2=n$. In our example we have

$$z=\tfrac{1}{2}\log_e\left\{\frac{398.985}{256.308}\right\}=0.221$$

and from Fisher's tables, entered with degrees of freedom $n_1=24$ and $n_2=26$, we find that z is less than the 5% point. We accept the hypothesis, therefore, and conclude that when the factor of chronological age is removed, the differences in mental age between pairs and within pairs of fraternal twins are of the same order of magnitude.

We may obtain two measures of the degree of relationship between mental age and chronological age from the results shown in Table XXXVIII. From the first row, for between pairs, we have

$$r_1 = \frac{9693.385}{\sqrt{(16162.231)(15389.308)}} = 0.615$$

and from the last row, for the total we have

$$r_2 = \frac{9693.385}{\sqrt{(16162.231)(22053.308)}} = 0.513$$

The first, $r=0.615$, is the better measure of the degree of relationship; in the second the relationship is masked by the inclusion of the within pairs differences in mental age.

***Example* 10.** *Relationship Between the Scores of Pupils on Two Mental Tests.*

In measuring the relationship between the scores of pupils on two mental tests, we frequently meet with the problem of deciding whether or not it is permissible to combine the results from several classes or grades. There is an additional problem here which is not generally recognized: that of determining the effect of combining the results on our measure of the relationship between the variables. As these problems are related, we shall consider both of them in this example and apply the methods to the results for five classes in Grade X. Part of the data has already been analyzed in Example 7; we found that the classes were of equal variability but differed significantly in ability. In the following analysis we assume that the classes are of equal variability; this assumption is satisfied in our case but, of course, if it is not satisfied we cannot apply the methods developed here.

If we denote by X_{st} and Y_{st} the scores of the t-th individual in the s-th class on the first and second test, respectively, then we may write

$$X_{st} = A + B_s + z_{st} \tag{119}$$

$$Y_{st} = C + D_s + u_{st} \tag{120}$$

where $s=1, 2, \ldots, k$; $t=1, 2, \ldots, n_s$; k denotes the number of classes and n_s the number of individuals in the s-th class. For the first and second test, respectively, A and C are measures of the ability of all the pupils in the grade and are defined as the mean scores for all individuals and classes; B_s and D_s are measures of the ability of the s-th class; z_{st} and u_{st} are measures of the differences between individuals within classes, and are assumed to be normally distributed

about zero with constant standard deviations, σ_1 and σ_2, say. It follows from the definition of A and C that

$$\left.\begin{aligned}\sum_s B_s=0\\ \sum_s D_s=0\end{aligned}\right\} \tag{121}$$

We will consider the relationship of the scores on the second test to those on the first, i.e. the regression of Y_{st} on X_{st}; the same method may be used, of course, in investigating the relationship of the scores on the first test to those on the second. The relationship is actually more complex than it appears as we must distinguish between two relationships: (1) the relationship between the estimates of the ability of the classes, and (2) the relationship between the estimates of the abilities of the individuals within each of the classes. If the relationship within classes is the same for each class, then we have only the relationships between and within classes, but we must first determine whether there is a common within classes relationship.

Assuming that the relationships are linear in form, we may write

$$Y_{st}=a+bB_s+c_s z_{st}+S_{st} \tag{122}$$

or, using the estimates of B_s and z_{st} from our previous analysis,

$$Y_{st}=a+b(\overline{X}_{s\cdot}-\overline{X}_{\cdot\cdot})+c_s(X_{st}-\overline{X}_{s\cdot})+S_{st} \tag{123}$$

where a, b and c_s are constants to be determined from the data; b and c_s are the regression coefficients for between and within classes, respectively, and S_{st} is a measure of the errors of estimating the scores on the second test from those on the first. We assume, further, that the quantity S_{st} is normally distributed about zero with constant standard deviation σ, constant for all classes. One of the first tests we shall have to make is a test of the validity of this assumption. If it is not satisfied, then we shall not have a common standard deviation, σ, but a standard deviation, σ_s say, for each class. It is necessary, therefore, to test the hypothesis

$$H_1\colon \sigma_s=\sigma \tag{124}$$

before we proceed with the analysis.

For each s, equation (123) may be written in the form

$$Y_{st}=d+c_s(X_{st}-\overline{X}_{s\cdot})+S_{st} \tag{125}$$

where d is a constant to be determined from the data. Denote by

$$\chi_s^2=\sum_t\{Y_{st}-d-c_s(X_{st}-\overline{X}_{s\cdot})\}^2 \tag{126}$$

and minimize χ_s^2 with regard to d and c_s to obtain the minimum value of χ_s^2, which we may denote by θ_s, as our estimate of $f_s\sigma_s^2$. We find

$$\theta_s = \sum_t Y_{st}^2 - \frac{(\sum_t Y_{st})^2}{n_s} - \frac{\left[\sum_t \{X_{st}Y_{st}\} - \frac{(\sum_t X_{st})(\sum_t Y_{st})}{n_s}\right]^2}{\sum_t X_{st}^2 - \frac{(\sum_t X_{st})^2}{n_s}} \tag{127}$$

To test the hypothesis H_1, we proceed as in Examples 4 and 7, following the method developed by Welch (**13**). The original data, in summary form, are given in the following table.

TABLE XL
DATA RELATING TO THE SCORES OF PUPILS ON THE TWO MENTAL TESTS

Class	Number of Pupils	Sum of Scores on First Test	Sum of Scores on Second Test	Sum of Squares of Scores on First Test	Sum of Squares of Scores on Second Test	Sum of Products of Scores on First and Second Test
s	n_s	$\sum_t X_{st}$	$\sum_t Y_{st}$	$\sum_t X_{st}^2$	$\sum_t Y_{st}^2$	$\sum_t Y_{st}X_{st}$
1	33	928	3,753	28,030	439,065	108,513
2	28	760	3,122	22,750	357,918	87,923
3	31	1,013	3,782	35,287	478,686	128,069
4	34	1,335	4,786	56,637	694,860	195,525
5	31	748	3,366	21,336	383,646	86,685
Total	157	4,784	18,809	164,040	2,354,175	606,715

From these data we calculate the quantity θ_s for each class, the value of L_1 for the test of the hypothesis H_1, and refer to Nayer's tables with $k=5$ and degrees of freedom $\bar{f}=29.4$. The values of θ_s, and the other quantities required in the calculation of L_1 are shown in Table XLI.

TABLE XLI

VALUES OF θ_s, AND CALCULATION OF L_1

Class	n_s	$\log n_s$	$n_s \log n_s$	θ_s	$\log \theta_s$	$n_s \log \theta_s$
1	33	1.51851	—	7,671.8638	3.88490	—
2	28	1.44716	—	5,039.2140	3.70236	—
3	31	1.49136	—	8,083.2055	3.90758	—
4	34	1.53148	—	7,453.5886	3.87237	—
5	31	1.49136	—	9,073.1429	3.95776	—
Total	157	$\sum_s n_s \log n_s = 235.16609$		37,321.0148	$\sum_s n_s \log \theta_s = 607.35390$	

We obtain $L_1 = 0.988$, and from Nayer's tables we find that L_1 is greater than the 5% point, so we accept the hypothesis H_1. The assumption of a constant standard deviation is valid, therefore, and we may proceed with the analysis.

Denote by

$$\chi^2 = \sum_s \sum_t \{ Y_{st} - a - b(\bar{X}_{s\cdot} - \bar{X}_{\cdot\cdot}) - c_s(X_{st} - \bar{X}_{s\cdot}) \}^2 \tag{128}$$

and minimize χ^2 with regard to all the parameters a, b and c_s to obtain the absolute minimum value of χ^2, χ_a^2.

We have

$$a = \frac{1}{N} \sum_s \sum_t Y_{st} = \bar{Y}_{\cdot\cdot}\,; \quad N = \sum_s n_s \tag{129}$$

$$b = \frac{\sum_s \left[\frac{(\sum_t Y_{st})(\sum_t X_{st})}{n_s} \right] - \frac{(\sum_s \sum_t Y_{st})(\sum_s \sum_t X_{st})}{N}}{\sum_s \left[\frac{(\sum_t X_{st})^2}{n_s} \right] - \frac{[\sum_s \sum_t X_{st}]^2}{N}} \tag{130}$$

$$c_s = \frac{\sum_t \{ Y_{st} X_{st} \} - \frac{(\sum_t X_{st})(\sum_t Y_{st})}{n_s}}{\sum_t X_{st}^2 - \frac{(\sum_t X_{st})^2}{n_s}} \tag{131}$$

for $s = 1, 2, \ldots, k$.

$$\chi_a^2=\sum_s\sum_t Y_{st}^2-\frac{(\sum_s\sum_t Y_{st})^2}{N}$$

$$-\frac{\left\{\sum_s\left[\frac{(\sum_t Y_{st})(\sum_t X_{st})}{n_s}\right]-\frac{(\sum_s\sum_t Y_{st})(\sum_s\sum_t X_{st})}{N}\right\}^2}{\sum_s\left[\frac{(\sum_t X_{st})^2}{n_s}\right]-\frac{[\sum_s\sum_t X_{st}]^2}{N}}$$

$$-\sum_s\left\{\frac{\left[\sum_t\{X_{st}Y_{st}\}-\frac{(\sum_t X_{st})(\sum_t Y_{st})}{n_s}\right]^2}{\sum_t X_{st}^2-\frac{(\sum_t X_{st})^2}{n_s}}\right\} \tag{132}$$

If there is a common within classes relationship, then $c_s=c$ and we may write

$$Y_{st}=a+b(\overline{X}_{s\cdot}-\overline{X}_{\cdot\cdot})+c(X_{st}-\overline{X}_{s\cdot})+S_{st} \tag{133}$$

The hypothesis we shall have to test before using the form shown in (133) is

$$H_2\colon c_s=c \tag{134}$$

This test may be developed in a manner similar to that used in the other examples. We minimize χ^2 subject to the condition that H_2 is true, i.e. we minimize

$$\chi^2=\sum_s\sum_t\{Y_{st}-a-b(\overline{X}_{s\cdot}-\overline{X}_{\cdot\cdot})-c(X_{st}-\overline{X}_{s\cdot})\}^2 \tag{135}$$

with regard to the parameters a, b and c, to obtain the relative minimum value of χ^2, χ_r^2, say. The estimates of a and b are, in this case, the same as those given in equations (129) and (130), respectively. The other values are as follows:

$$c=\frac{\sum_s\left[\sum_t Y_{st}X_{st}-\frac{(\sum_t Y_{st})(\sum_t X_{st})}{n_s}\right]}{\sum_s\left[\sum_t X_{st}^2-\frac{(\sum_t X_{st})^2}{n_s}\right]}. \tag{136}$$

$$\chi_r^2=\chi_a^2+\chi_1^2,\quad \text{say} \tag{137}$$

$$\text{where } \chi_1^2 = \sum_s \left\{ \frac{\left[\sum_t \{X_{st}Y_{st}\} - \frac{(\sum_t Y_{st})(\sum_t X_{st})}{n_s} \right]^2}{\sum_t X_{st}^2 - \frac{(\sum_t X_{st})^2}{n_s}} \right\}$$

$$- \frac{\left[\sum_s \left\{ \sum_t (X_{st}Y_{st}) - \frac{(\sum_t X_{st})(\sum_t Y_{st})}{n_s} \right\} \right]^2}{\sum_s \left[\sum_t X_{st}^2 - \frac{(\sum_t X_{st})^2}{n_s} \right]} \qquad (138)$$

To test the hypotheses $H_2 : c_s = c$, we calculate

$$z = \tfrac{1}{2} \log_e \left\{ \frac{\chi_1^2}{n_1} \bigg/ \frac{\chi_a^2}{n_2} \right\} \qquad (139)$$

and refer to Fisher's tables of z with degrees of freedom $n_1 = k-1$ and $n_2 = N-k-2$. In the evaluation of χ_a^2 and χ_1^2, we need the quantities shown in the following table (calculated from the data in Table XL).

TABLE XLII

Quantities Required in the Evaluation of χ_a^2 and χ_1^2

Class	$\frac{(\sum_t Y_{st})(\sum_t X_{st})}{n_s}$	$\frac{(\sum_t X_{st})^2}{n_s}$	$\sum_t X_{st}^2 - \frac{(\sum_t X_{st})^2}{n_s}$	$\sum_t X_{st}Y_{st} - \frac{(\sum_t X_{st})(\sum_t Y_{st})}{n_s}$
1	105,538.9091	26,096.4848	1,933.5152	2,974.0909
2	84,740.0000	20,628.5714	2,121.4286	3,183.0000
3	123,586.0000	33,102.2258	2,184.7742	4,483.0000
4	187,920.8824	52,418.3824	4,218.6176	7,604.1176
5	81,218.3226	18,048.5161	3,287.4839	5,466.6774
Total	583,004.1141	150,294.1805	13,745.8195	23,710.8859

Substituting these values in equations (132) and (138), we find

$$\chi_a^2 = 37{,}912.5059$$
$$\chi_1^2 = 446.0445$$

and for substitution in equation (139) we calculate

$$\frac{\chi_1^2}{n_1} = \frac{446.0445}{4} = 111.5111$$

$$\frac{\chi_a^2}{n_2} = \frac{37912.5059}{150} = 252.7500$$

Since χ_1^2/n_1 is actually less than χ_a^2/n_2, it is clear that the departures from the common relationship are not significant. We may accept the hypothesis H_2, therefore, and conclude that there is a common within classes relationship or regression.

Since the assumption $c_s = c$ is valid, we may present the final results of our analysis in the form shown below; this is the form generally used for analyses of this type.

TABLE XLIII

ANALYSIS OF VARIANCE AND COVARIANCE OF SCORES OF PUPILS ON THE TWO MENTAL TESTS

Variance	d.f.	Sum of Squares Second Test	Sum of Squares First Test	Sum of Products
Between Classes . .	4	22,141.6675	4,519.3015	9,868.7256
Within Classes . . .	152	78,667.2115	13,745.8195	23,710.8859
Total	156	100,808.8790	18,265.1210	33,579.6115

From the results we obtain the values shown in the following table.

TABLE XLIV

REGRESSION COEFFICIENTS, CORRELATION COEFFICIENTS AND ADJUSTED SUM OF SQUARES FOR SECOND TEST

Factor	Regression Coefficient	Correlation Coefficient	Adjusted Analysis for Second Test		
			d.f.	Sum of Squares	Mean Square
Between Classes	2.184	0.987	3	591.4911	197.164
Within Classes	1.725	0.721	151	37,767.0593	250.113
Total	1.838	0.783	154	38,358.5504	——

From the adjusted analysis we find, as is to be expected, that the differences between the classes on the second test are not significant when corrected for the differences found on the first.

From the values of the correlation coefficients, we see that the relationships for between and within classes are very different; this also appears, of course, when we consider the regression coefficients. The relationships have been shown graphically in Figure 1; the regression lines shown are those determined by the regression coefficients given in Table XLV.

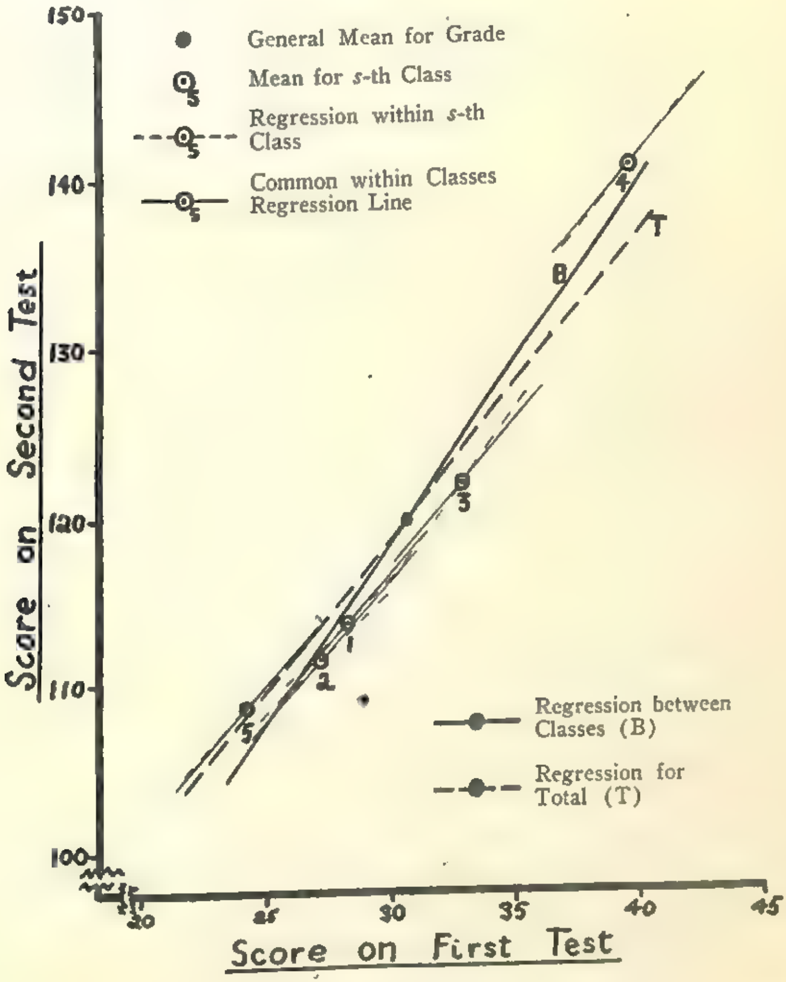

FIGURE 1

Relationship between Scores of Pupils on two Mental Tests

TABLE XLV

REGRESSION COEFFICIENTS

Factor	Regression Coefficients
Between Classes	$b = 2.184$
Within Classes	$c_1 = 1.538$
	$c_2 = 1.500$
	$c_3 = 2.052$
	$c_4 = 1.803$
	$c_5 = 1.663$
	$c = 1.725$
Total	$c^1 = 1.838$

It is clear that the three general relationships: (1) between means of classes, (2) within classes, and (3) total for all classes, are very different. The choice of a particular one for use will be determined mainly by the nature of the problem considered; for instance, if we are interested in the relationship within classes we shall choose the second one, $c = 1.725$, and ignore the others. The usefulness of an analysis of this kind lies in the fact that it shows clearly the different relationships, and, at the same time, enables us to choose the appropriate measure for use in the solution of our educational problem.

Note:

In some cases we may wish to determine whether our measure of the relationship, the regression coefficient, is significantly greater than zero. As this problem is related to those discussed above, we shall consider here the test of the corresponding hypothesis for a particular class, the first, say. The hypothesis we wish to test, therefore, may be written

$$H_0: c_1 = 0 \tag{140}$$

The test of H_0 is developed by the usual method. We denote by

$$\chi^2 = \sum_i \{y_i - d - c_1(X_i - \overline{X}.)\}^2 \tag{141}$$

(see equation (126)) and minimize χ^2:

(1) with regard to d and c_1 to obtain the absolute minimum value of χ^2, χ_a^2;

(2) with regard to the parameters remaining if H_0 is true, i.e. if $c_1 = 0$, to obtain the relative minimum value of χ^2, χ_r^2.

We obtain

$$\chi_a^2 = \sum_t y_t^2 - \frac{(\sum_t y_t)^2}{n} - \frac{\left[\sum_t \{y_t X_t\} - \frac{(\sum_t X_t)(\sum_t y_t)}{n}\right]^2}{\sum_t X_t^2 - \frac{(\sum_t X_t)^2}{n}} \tag{142}$$

$$\chi_r^2 = \chi_a^2 + \chi_0^2, \quad \text{say} \tag{143}$$

where

$$\chi_0^2 = \frac{\left[\sum_t \{y_t X_t\} - \frac{(\sum_t y_t)(\sum_t X_t)}{n}\right]^2}{\sum_t X_t^2 - \frac{(\sum_t X_t)^2}{n}} \tag{144}$$

and n denotes the number of pupils in the class considered.

To test the hypothesis H_0, we calculate

$$z = \tfrac{1}{2} \log_e \left\{ \frac{\chi_0^2}{n_1} \Big/ \frac{\chi_a^2}{n_2} \right\} \tag{145}$$

and refer to Fisher's tables of z with degrees of freedom $n_1 = 1$ and $n_2 = n - 2$. For the first class we have

$$\chi_a^2 = 7671.8638$$
$$\chi_0^2 = 4574.6817$$
$$n = 33$$

$$z = \tfrac{1}{2} \log_e \left\{ \frac{4574.6817}{\frac{7671.8638}{31}} \right\} = 1.459$$

From Fisher's tables, entered with degrees of freedom $n_1 = 1$ and $n_2 = 31$, we find that z is greater than the 1% point. We reject the hypothesis H_0, and conclude that the regression coefficient is significantly greater than zero.

The tests of significance of the other regression coefficients considered above may be developed in a similar manner.

Example 11. *Reliability of Mental Tests: Relationship between the Sampling Unit and the Estimates of Reliability of the Test.*

In determining the reliability of a mental test, we meet with a difficulty similar to that considered in the previous example. The tests

are administered to groups of pupils, the school class is generally the unit chosen, and these groups form our sampling units. Estimates of the reliability of the test may be calculated separately for each class, or other such unit, and these results may be interpreted without difficulty as the statistical and sampling units are the same. We may, on the other hand, wish to use the same data in estimating the reliability of the test for a larger unit, say the grade. Since it will be necessary to combine the original data into larger groups, the statistical and sampling units will no longer be the same and the interpretation of the results may be difficult. In this example we shall consider the problems of determining: (1) the effect of combining the data on our estimates of reliability, (2) the conditions which must be satisfied before the results may be combined, and (3) the interpretation of the results. The data used in illustrating the methods refer to the scores of high school pupils on two forms of a mental test; the coefficients given below, therefore, are the Comparable Forms estimates of the reliability of the test.

The problem of determining which group—the class, grade or school as a whole—should be chosen as the unit is related to those stated above. It is, however, an educational rather than a statistical problem and will not be considered here. A more detailed study of this particular problem, using methods similar to those developed here, will be published later.

As the arguments are easier to follow when we have a particular case in mind, we shall begin with the analysis of the results for a group of Grade XI pupils. The original data, in summary form, for the four classes in this grade are given in the following table.

TABLE XLVI

SCORES OF GRADE XI PUPILS ON TWO FORMS OF A MENTAL TEST

Class (s)	Number of Pupils n_s	Sum of Scores		Sum of Squares of Scores		Sum of Products
		1st Form	2nd Form	1st Form	2nd Form	
1	32	1,180	1,352	48,092	60,402	53,265
2	21	632	782	20,436	31,030	24,984
3	41	1,084	1,292	32,062	44,064	36,969
4	21	495	671	13,465	24,161	17,634
Total	115	3,391	4,097	114,055	159,657	132,852

Let us consider first an analysis of variance and covariance of these scores similar to that considered in the previous example. Applying the L_1 criterion discussed in the other examples, we find that the classes are of equal variability; we may, therefore, present the results in the form shown below.

TABLE XLVII

ANALYSIS OF VARIANCE AND COVARIANCE OF SCORES OF PUPILS ON THE TWO FORMS

Variance	d.f.	Sum of Squares		Sum of Products	Correlation Coefficient r
		First Form	Second Form		
Between Classes .	3	2,870.1804	2,435.9160	2,557.0634	0.967
Within Classes .	111	11,194.5500	11,261.0058	9,486.8757	0.845
Total . . .	114	14,064.7304	13,696.9218	12,043.9391	0.868

The correlation coefficients shown in the last column of the table are the measures usually employed in estimating the reliability of the test. The difference between the reliability coefficients for "within classes", $r=0.845$, and the "total", $r=0.868$, is a measure of the effect of combining the data into a larger group. The latter coefficient refers to the reliability of the test for the grade, and is larger because of the inclusion of the very significant differences between classes.

This analysis, therefore, gives us a measure of the effect of combining the results. We can see more clearly what is happening, however, if we consider analyses of the type shown in Example 2. The results for each class, presented in a form similar to that shown in Tables V and VII, are given in Table XLVIII. Underlying these analyses is the assumption that we may write

$$X_{sti}=A+B_s+D_{si}+C_{st}+z_{sti} \tag{146}$$

where X_{sti} denotes the score of the t-th individual in the s-th class on the i-th form of the test; $i=1, 2$; $s=1, 2, \ldots, k$; $t=1, 2, \ldots, n_s$; k denotes the number of classes and n_s the number of pupils in the s-th class. A is a measure of the common ability of all individuals in the grade and is defined as the mean of the scores for all individuals, forms and classes; B_s is a measure of the ability of the s-th class; D_{si} is a measure of the difference between the scores on the two forms of the test, i.e. the practice effect, for the s-th class; C_{st} is a measure of the

TABLE XLVIII

ANALYSIS OF VARIANCE OF SCORES OF PUPILS ON THE TWO FORMS OF THE MENTAL TEST (BY CLASSES)

Variance	Class 1			Class 2			Class 3			Class 4		
	d.f.	Sum of Squares	Mean Square	d.f.	Sum of Squares	Mean Square	d.f.	Sum of Squares	Mean Square	d.f.	Sum of Squares	Mean Square
Between Forms. . . .	1	462.2500	462.2500	1	535.7143	535.7143	1	527.6097	527.6097	1	737.5238	737.5238
Between Individuals. .	31	7,339.7500	236.7661	20	3,112.3333	155.6167	40	6,185.9513	154.6488	20	4,076.6190	203.8310
Error	31	519.7500	16.7661	20	213.2857	10.6643	40	566.3902	14.1598	20	441.4762	22.0738
Total	63	8,321.7500	—	41	3,861.3333	—	81	7,279.9512	—	41	5,255.6190	—

ability of the t-th individual in the s-th class; z_{sti} represents the errors of measurement. It follows from the definition of A that

$$\left.\begin{aligned} \sum_s B_s &= 0 \\ \sum_i D_{si} &= 0 \\ \sum_t C_{st} &= 0 \end{aligned}\right\} \tag{147}$$

In combining the results, we also assume that

(1) z_{sti} is normally distributed about zero with constant standard deviation, σ say, constant for all classes ;

(2) C_{st} is normally distributed about zero with constant standard deviation, σ_c say, constant for all classes.

If these assumptions are not satisfied, then, instead of σ and σ_c, we shall have σ_s and σ_{cs}, respectively, i.e. standard deviations constant for each class only. We must, therefore, test the hypotheses

$$H_1 : \sigma_s = \sigma \tag{148}$$

and

$$H_2 : \begin{cases} \sigma_{cs} = \sigma_c \\ \sigma_s \ = \sigma \end{cases} \tag{149}$$

(see equations (64) and (68) of Example 4) before combining the results.

The tests of these hypotheses are the same as those discussed in Example 4 where the problem was similar. The relevant data and the method to be followed in evaluating $\log L_1$, where L_1 is the criterion used in testing these hypotheses, are shown in Tables XLIX and L. For the test of the hypothesis H_1, we have $L_1 = 0.974$ and from Nayer's tables, entered with $k = 4$ and mean degrees of freedom $\bar{f} = 27.8$, we find that L_1 is greater than the 5% point. We may accept the hypothesis H_1, and proceed to the test of the hypothesis $H_2 : \begin{cases} \sigma_{cs} = \sigma_c \\ \sigma_s \ = \sigma \end{cases}$. For the test of this hypothesis, we have $L_1 = 0.983$ and referring to Nayer's tables with $k = 4$ and mean degrees of freedom $\bar{f} = 55.5$ we find that L_1 is greater than the 5% point. We may, therefore, accept both H_1 and H_2 and, since our assumptions are satisfied, we may combine the results.

TABLE XLIX

TEST OF THE HYPOYHESIS H_1: $\sigma_s = \sigma$; EVALUATION OF LOG L_1

n_s	$\log n_s$	$n_s \log n_s$	θ'_s	$\log \theta'_s$	$n_s \log \theta'_s$
64	1.80618	——	519.7500	2.71579	——
42	1.62325	——	213.2857	2.32896	——
82	1.91381	——	566.3902	2.75312	——
42	1.62325	——	441.4762	2.64491	——
230	$\sum_s n_s \log n_s = 408.88094$		1,740.9021	$\sum_s n_s \log \theta'_s = 608.46894$	

$$\log L_1 = \bar{1}.98874$$

TABLE L

TEST OF THE HYPOTHESIS H_2: $\begin{cases} \sigma_{cs} = \sigma_c \\ \sigma_s = \sigma \end{cases}$; EVALUATION OF LOG L_1

n_s	$\log n_s$	$n_s \log n_s$	θ_s	$\log \theta_s$	$n_s \log \theta_s$
64	1.80618	——	7,859.5000	3.89539	——
42	1.62325	——	3,325.6190	3.52187	——
82	1.91381	——	6,752.3415	3.82945	——
42	1.62325	——	4,518.0952	3.65495	——
230	$\sum_s n_s \log n_s = 408.88094$		22,455.5557	$\sum_s n_s \log \theta_s = 864.74630$	

$$\log L = \bar{1}.99244$$

Although it is not necessary to do so at this stage, it is convenient to consider the test of another hypothesis before we combine the results. It will be seen that in equation (146), and subsequent discussion, we have considered the practice effect, D_{si}, separately for each class. It is possible that the practice effect is constant for all classes, i.e. that $D_{si} = D_i$, say, for all s. We may develop a test of this hypothesis

$$H_3': D_{si} = D_i \tag{150}$$

by the usual method. Denote by

$$\chi^2 = \sum_s \sum_t \sum_i \{X_{sti} - A - B_s - D_{si} - C_{st}\}^2 \tag{151}$$

and minimize χ^2: (1) with regard to all the parameters in (151) to obtain the absolute minimum value of χ^2, χ_a^2; and (2) with regard to

the parameters remaining if H_3 is true, to obtain the relative minimum value of χ^2, χ_r^2.

We obtain

$$\chi_a^2 = \sum_s \left[\tfrac{1}{2} \sum_t (X_{st1} - X_{st2})^2 - \frac{\left\{ \sum_t (X_{st1} - X_{st2}) \right\}^2}{2n_s} \right] \tag{152}$$

(see equation (46) of Example 2)

$$\chi_r^2 = \chi_a^2 + \chi_1^2, \quad \text{say} \tag{153}$$

where

$$\chi_1^2 = \sum_s \left[\frac{\sum_i (\sum_t X_{sti})^2}{n_s} - \frac{(\sum_t \sum_i X_{sti})^2}{2n_s} \right] - \left[\frac{\sum_i \{ \sum_s \sum_t X_{sti} \}^2}{\sum_s n_s} - \frac{\{ \sum_s \sum_t \sum_i X_{sti} \}^2}{2 \sum_s n_s} \right] \tag{154}$$

To test the hypothesis H_3, we calculate

$$z = \tfrac{1}{2} \log_e \left\{ \frac{\chi_1^2}{n_1} \middle/ \frac{\chi_a^2}{n_2} \right\} \tag{155}$$

and refer to Fisher's tables of z with degrees of freedom $n_1 = k - 1$ and $n_2 = \sum_s (n_s - 1)$. The results of the analysis may be shown as in the following table:

TABLE LI

TEST OF H_3: $D_{ct} = D_t$, ANALYSIS OF VARIANCE OF COMBINED RESULTS

Variance	d.f.	Sum of Squares	Mean Square
Due to Common Practice Effect	1	2,167.1130	2,167.1130
Due to Departures from Common Practice Effect	3	95.9848	31.9949
Between Means of Classes	3	5,210.1117	1,736.7039
Between Individuals (Within Classes)	111	20,714.6536	186.6185
Error (Within Classes)	111	1,740.9021	15.6838
Total	229	29,928.7652	——

We have

$$\chi_1^2 = 95.9848$$
$$\chi_a^2 = [illegible]740.9021$$
$$= [illegible]11$$

Substituting these values in (155), we obtain

$$z = \tfrac{1}{2} \log_e \left\{\frac{31.9949}{15.6838}\right\} = 0.356$$

and from Fisher's tables, entered with degrees of freedom $n_1 = 3$ and $n_2 = 111$, we find that z is less than the 5% point. We accept the hypothesis H_3, therefore, and conclude that the practice effect is constant for all classes. Our final analysis will be as shown in Table LII.

TABLE LII

FINAL ANALYSIS OF VARIANCE OF COMBINED RESULTS

Variance	d.f.	Sum of Squares	Mean Square
Due to Common Practice Effect	1	2,167.1130	2,167.1130
Between Means of Classes	3	5,210.1117	1,736.7039
Between Individuals (Within Classes) . .	111	20,714.6536	186.6185
Error (Within Classes)	114	1,836.8869	16.1130
Total	229	29,928.7652	———

If we had ignored the differences between the classes and analyzed the results for the grade as a whole, we would have found the results given in the following table:

TABLE LIII

ANALYSIS OF VARIANCE OF SCORE OF PUPILS; FOR THE GRADE AS A UNIT

Variance	d.f.	Sum of Squares	Mean Square
Between Forms	1	2,167.1130	2,167.1130
Between Individuals	114	25,924.7653	227.4102
Error (Within Classes)	114	1,836.8869	16.1130
Total	229	29,928.7652	———

Comparing the results of this analysis with those given in Table LII, we find that there is only one difference. The "between individuals" sum of squares of Table LIII is broken up into two parts in Table LII—one ascribable to differences between means of classes and the other to

differences between individuals within classes. In conditions similar to those existing here, the effect of using the larger unit is exactly this merging of the two factors. The analysis shown in Table LII, therefore, gives us a measure of the effect of combining the results into a larger unit.

If the practice effect is not constant, but the other conditions are satisfied, then the analysis shown in Table LI will give us the desired measure of the effect of combining the results. In this case, of course, both the error and the between individuals variance in the combined analysis of the scores of all pupils in the grade, as in Table LIII, will be increased.

Let us consider, finally, the interpretation of the results for a case where the conditions are not satisfied. The data given below in summary form refer to the scores of pupils on the two forms of the mental test considered above, but the unit in this case is the grade, not the class.

TABLE LIV

SCORES OF PUPILS ON TWO FORMS OF A MENTAL TEST (BY GRADES)

Grade	Number of Pupils n_s	Sum of Scores		Sum of Squares of Scores		Sum of Products
		First Form	Second Form	First Form	Second Form	
X	157	4,062	4,784	119,544	164,040	137,593
XI	115	3,391	4,097	114,055	159,657	132,852
XII	115	4,484	5,025	194,382	237,933	213,023
XIII	41	1,858	2,095	91,910	113,661	101,311
Total	428	13,795	16,001	519,891	675,291	584,779

The analyses by grades are given in Table LV; in these analyses, of course, we ignore the effect of combining the results for all classes within a grade. As in the previous case, we test the hypotheses H_1 and H_2 before considering the final combination of the results for the school as a whole. From the values given in Tables LVI and LVII, we find that:

(1) For the test of the hypothesis $H_1 : \sigma_s = \sigma$, we obtain $L_1 = 0.999$ and from Nayer's tables, entered with $k = 4$ and mean degrees of freedom $\bar{f} = 106$, we find that L_1 is greater than the 5% point. We may accept the hypothesis H_1 and, as far as the error term is concerned, combine the results.

TABLE LV

ANALYSIS OF VARIANCE OF SCORES OF PUPILS ON THE TWO FORMS OF THE MENTAL TEST (BY GRADES)

Variance	Grade X			Grade XI			Grade XII			Grade XIII		
	d.f.	Sum of Squares	Mean Square	d.f.	Sum of Squares	Mean Square	d.f.	Sum of Squares	Mean Square	d.f.	Sum of Squares	Mean Square
Between Forms . .	1	1,660.1402	1,660.1402	1	2,167.1130	2,167.1130	1	1,272.5261	1,272.5261	1	684.9878	684.9878
Between Individuals	156	30,175.7134	193.4341	114	25,924.7653	227.4102	114	36,045.3653	316.1874	40	13,532.9756	338.3244
Error	156	2,538.8598	16.2747	114	1,836.8869	16.1130	114	1,861.9738	16.3331	40	789.5122	19.7378
Total	313	34,374.7134	—	229	29,928.7652	—	229	39,179.8652	—	81	15,007.4756	—

TABLE LVI

TEST OF THE HYPOTHESIS H_1: $\sigma_s = \sigma$; EVALUATION OF LOG L_1

n_s	$\log n_s$	$n_s \log n_s$	θ'_s	$\log \theta'_s$	$n_s \log \theta'_s$
314	2.49693	——	2,538.8598	3.40464	——
230	2.36173	——	1,836.8869	3.26408	——
230	2.36173	——	1,861.9738	3.26997	——
82	1.91381	——	789.5122	2.89736	——
856	$\sum_s n_s \log n_s = 2{,}027.36424$		7 027.2327	$\sum_s n_s \log \theta'_s = 2{,}809.47198$	

$\log L_1 = \bar{1}.99936$

TABLE LVII

TEST OF THE HYPOTHESIS H_2: $\begin{cases} \sigma_{cs} = \sigma_c \\ \sigma_s = \sigma \end{cases}$; EVALUATION OF LOG L_1

n_s	$\log n_s$	$n_s \log n_s$	θ_s	$\log \theta_s$	$n_s \log \theta_s$
314	2.49693	——	32,714.5732	4.51474	——
230	2.36173	——	27,761.6522	4.44345	——
230	2.36173	——	37,907.3391	4.57872	——
82	1.91381	——	14,322.4878	4.15602	——
856	$\sum_s n_s \log n_s = 2{,}027.36424$		112,706.0523	$\sum_s n_s \log \theta_s = 3{,}833.52110$	

$\log L_1 = \bar{1}.99051$

(2) For the test of the hypothesis H_2: $\begin{cases} \sigma_{cs} = \sigma_c \\ \sigma_s = \sigma \end{cases}$, we obtain $L_1 = 0.978$ and from Nayer's tables, entered with $k = 4$ and mean degrees of freedom $\bar{f} = 212$, we find that L_1 is less than the 1% point. We reject the hypothesis H_2, therefore, and conclude that the standard deviation, σ_{cs}, varies from grade to grade. Our combined analysis, for the test of the hypothesis H_3: $D_{si} = D_i$, will be in the form shown in Table LVIII.

TABLE LVIII

TEST OF THE HYPOTHESIS H_3: $D_{si}=D_i$; ANALYSIS OF VARIANCE OF COMBINED RESULTS

Variance		d.f.	Sum of Squares	Mean Square
Due to Common Practice Effect		1	5,685.0887	5,685.0887
Due to Departures from Common Practice Effect		3	99.6784	33.2261
Between Means of Grades		3	39,539.7599	13,179.9200
Between Individuals in	Grade X	156	30,175.7134	193.4341
	Grade XI	114	25,924.7653	227.4102
	Grade XII	114	36,045.3653	316.1874
	Grade XIII	40	13,532.9756	338.3244
Error (Within Grades)		424	7,027.2327	16.5737
Total		855	158,030.5793	———

For the test of the hypothesis H_3: $D_{si}=D_i$, we have

$$\chi_a^2=7027.2328$$
$$\chi_1^2=\quad 99.6784$$
$$n_1=3;\ n_2=424$$

and we find

$$z=\tfrac{1}{2}\log_e\left\{\frac{33.2261}{16.5737}\right\}=0.348$$

From Fisher's tables, entered with degrees of freedom $n_1=3$ and $n_2=424$, we find that z is less than the 5% point. We accept the hypothesis H_3: $D_{si}=D_i$, and conclude that the departures from the common practice effect are not significant. We may, therefore, present our final analysis in the form shown in Table LIX.

TABLE LIX

FINAL ANALYSIS OF VARIANCE OF THE COMBINED RESULTS (ALL GRADES)

Variance		d.f.	Sum of Squares	Mean Square
Due to Common Practice Effect		1	5,685.0887	5,685.0887
Between Means of Grades		3	39,539.7599	13,179.9200
Between Individuals in	Grade X	156	30,175.7134	193.4341
	Grade XI	114	25,924.7653	227.4102
	Grade XII	114	36,045.3653	316.1874
	Grade XIII	40	13,532.9756	338.3244
Error (Within Grades)		427	7,126.9111	16.6907
Total		855	158,030.5793	———

The effect of ignoring the differences between grades and analyzing the results for the school as a whole is shown in Table LX.

TABLE LX

EFFECT OF COMBINING THE RESULTS FOR ALL GRADES ON THE ANALYSIS OF VARIANCE

Variance		d.f.	Sum of Squares	Mean Square
Between Means of Grades		3	39,539.7599	13,179.9200
Between Individuals in	Grade X	156	30,175.7134	193.4341
	Grade XI	114	25,924.7653	227.4102
	Grade XII	114	36,045.3653	316.1874
	Grade XIII	40	13,532.9756	338.3244
Total {Between Individuals for all Grades}		427	145,218.5795	340.0904

The differences between grades are shown in the first five rows; if we combine these results we assume that there is a common variance, which clearly is not the case, and use the mean square in the last row, 340.0904, as the estimate of it.

The interpretation of the results given in Table LIX is not difficult. We have a significant practice effect, significant differences between the ability of pupils in different grades, and significant differences between the variability of the scores of pupils in different grades. If

we assume that the discriminating power of a mental test may be measured by the degree to which it differentiates between the individuals of a particular group, it is clear that the test considered is more suitable for use in the higher grades. Finally, we find that the errors of measurement by means of the test are constant for all grades. The probable error of the score of any individual in the groups considered may be taken as

$$\text{P.E.}_{\text{Ind. Score}} = 0.6745\sqrt{16.6907}$$
$$= 2.8 \text{ score units.}$$

BIBLIOGRAPHY

1. Fisher, R. A. *Statistical Methods for Research Workers.* Edinburgh: Oliver & Boyd, 7th Edition, 1938. Pp. xv+356.
2. Jackson, R. W. B. "Reliability of Mental Tests". *British Journal of Psychology* (General Section), XXIX (Part 3, Jan. 1939), pp. 267-287.
3. Johnson, P. O. and Neyman, J. "Tests of Certain Linear Hypotheses and Their Application to Some Educational Problem". *Statistical Research Memoirs*, 1 (1936), Department of Applied Statistics, University College, London, W.C. 1., pp. 57-93.
4. Kolodziejczyk, St. "On an Important Class of Statistical Hypotheses". *Biometrika*, XXVII (1935), pp. 161.
5. Nayer, P. P. N. "An Investigation into the Application of Neyman and Pearson's L_1 Test, with Tables of Percentage Limits". *Statistical Research Memoirs*, 1 (1936), pp. 38-51.
6. Neyman, J. "Sur la Vérification des Hypothèses Statistiques Composées". *Bulletin de la Société mathématique.* France, LXIII (1935), pp. 246-266.
7. Neyman, J. and Pearson, E. S. "Contributions to the Theory of Testing Statistical Hypotheses: (1) Unbiassed Critical Regions of Type A and A_1". *Statistical Research Memoirs*, 1 (1936), pp. 1-37.
8. Neyman, J. and Pearson, E. S. "Sufficient Statistics and Uniformly Most Powerful Tests of Statistical Hypotheses". *Statistical Research Memoirs*, 1 (1936), pp. 113-137.
9. Neyman, J. and Pearson, E. S. "On the Problem of k Samples". *Bulletin de l'Académie Polonaise des Sciences et des Lettres*, Series A (1931), pp. 460-481.
10. Snedecor, George W. *Calculation and Interpretation of Analysis of Variance and Covariance.* Ames, Iowa: Collegiate Press, 1934. Pp. 96.
11. Snedecor, George W. *Statistical Methods.* Ames, Iowa: Collegiate Press, 1938. Pp. xiii+388.
12. Welch, B. L. "Some Problems in the Analysis of Regression Among k Samples of Two Variables". *Biometrika*, XXVII (1935), pp. 145-160.
13. Welch, B. L. "Note on an Extension of the L_1 Test". *Statistical Research Memoirs*, 1 (1936), pp. 52-56.
14. Wingfield, Alex. H. *Twins and Orphans, the Inheritance of Intelligence.* London: J. M. Dent & Sons Ltd., 1928. Pp. 127.

NOTE:—A discussion of the recent advances in the Analysis of Variance and Covariance, with a complete bibliography, may be found in the following paper: Cochran, W. G. "Recent Work on the Analysis of Variance" *Journal of the Royal Statistical Society.* Vol. CI (1938), Part II, pp. 434-449.

APPENDIX A

TABLE VI

5 Per Cent. Points of the Distribution of z

Values of n_2	Values of n_1: 1.	2.	3.	4.	5.	6.	8.	12.	24.	∞
1	2·5421	2·6479	2·6870	2·7071	2·7194	2·7276	2·7380	2·7484	2·7588	2·7693
2	1·4592	1·4722	1·4765	1·4787	1·4800	1·4808	1·4819	1·4830	1·4840	1·4851
3	1·1577	1·1284	1·1137	1·1051	1·0994	1·0953	1·0899	1·0842	1·0781	1·0716
4	1·0212	·9690	·9429	·9272	·9168	·9093	·8993	·8885	·8767	·8639
5	·9441	·8777	·8441	·8236	·8097	·7997	·7862	·7714	·7550	·7368
6	·8948	·8188	·7798	·7558	·7394	·7274	·7112	·6931	·6729	·6499
7	·8606	·7777	·7347	·7080	·6896	·6761	·6576	·6369	·6134	·5862
8	·8355	·7475	·7014	·6725	·6525	·6378	·6175	·5945	·5682	·5371
9	·8163	·7242	·6757	·6450	·6238	·6080	·5862	·5613	·5324	·4979
10	·8012	·7058	·6553	·6232	·6009	·5843	·5611	·5346	·5035	·4657
11	·7889	·6909	·6387	·6055	·5822	·5648	·5406	·5126	·4795	·4387
12	·7788	·6786	·6250	·5907	·5666	·5487	·5234	·4941	·4592	·4156
13	·7703	·6682	·6134	·5783	·5535	·5350	·5089	·4785	·4419	·3957
14	·7630	·6594	·6036	·5677	·5423	·5233	·4964	·4649	·4269	·3782
15	·7568	·6518	·5950	·5585	·5326	·5131	·4855	·4532	·4138	·3628
16	·7514	·6451	·5876	·5505	·5241	·5042	·4760	·4428	·4022	·3490
17	·7466	·6393	·5811	·5434	·5166	·4964	·4676	·4337	·3919	·3366
18	·7424	·6341	5753	·5371	·5099	·4894	·4602	·4255	·3827	·3253
19	·7386	·6295	·5701	·5315	·5040	·4832	·4535	·4182	·3743	·3151
20	·7352	·6254	·5654	·5265	·4986	·4776	·4474	·4116	·3668	·3057
21	·7322	·6216	·5612	·5219	·4938	·4725	·4420	·4055	·3599	·2971
22	·7294	·6182	·5574	·5178	·4894	·4679	·4370	·4001	·3536	·2892
23	·7269	·6151	·5540	·5140	·4854	·4636	·4325	·3950	·3478	·2818
24	·7246	·6123	·5508	·5106	·4817	·4598	·4283	·3904	·3425	·2749
25	·7225	·6097	·5478	·5074	·4783	·4562	·4244	·3862	·3376	·2685
26	·7205	·6073	·5451	·5045	·4752	·4529	·4209	·3823	·3330	·2625
27	·7187	·6051	·5427	·5017	·4723	·4499	·4176	·3786	·3287	·2569
28	·7171	·6030	·5403	·4992	·4696	·4471	·4146	·3752	·3248	·2516
29	·7155	·6011	·5382	·4969	·4671	·4444	·4117	·3720	·3211	·2466
30	·7141	·5994	·5362	·4947	·4648	·4420	·4090	·3691	·3176	·2419
60	·6933	·5738	·5073	·4632	·4311	·4064	·3702	·3255	·2654	·1644
∞	·6729	·5486	·4787	·4319	·3974	·3706	·3309	·2804	·2085	0

This table is reproduced by permission from "Statistical Methods for Research Workers" by Professor R. A. Fisher, published at 15/- by Oliver & Boyd, Edinburgh, and attention is drawn to the larger collection in "Statistical Tables" by Professor R. A. Fisher and F. Yates, published at 12/6 by Oliver & Boyd, Edinburgh.

TABLE VI.—*Continued*

1 Per Cent. Points of the Distribution of z

Values of n_2	Values of n_1: 1.	2.	3.	4.	5.	6.	8.	12	24.	∞.
1	4·1535	4·2585	4·2974	4·3175	4·3297	4·3379	4·3482	4·3585	4·3689	4·3794
2	2·2950	2·2976	2·2984	2·2988	2·2991	2·2992	2·2994	2·2997	2·2999	2·3001
3	1·7649	1·7140	1·6915	1·6786	1·6703	1·6645	1·6569	1·6489	1·6404	1·6314
4	1·5270	1·4452	1·4075	1·3856	1·3711	1·3609	1·3473	1·3327	1·3170	1·3000
5	1·3943	1·2929	1·2449	1·2164	1·1974	1·1838	1·1644	1·1457	1·1239	1·0997
6	1·3103	1·1955	1·1401	1·1068	1·0843	1·0680	1·0460	1·0218	·9948	·9643
7	1·2526	1·1281	1·0672	1·0300	1·0048	·9864	·9614	·9335	·9020	·8658
8	1·2106	1·0787	1·0135	·9734	·9459	·9259	·8983	·8673	·8319	·7904
9	1·1786	1·0411	·9724	·9299	·9006	·8791	·8494	·8157	·7769	·7305
10	1·1535	1·0114	·9399	·8951	·8646	·8419	·8104	·7744	·7324	·6816
11	1·1333	·9874	·9136	·8674	·8354	·8116	·7785	·7405	·6958	·6408
12	1·1166	·9677	·8919	·8443	·8111	·7864	·7520	·7122	·6649	·6061
13	1·1027	·9511	·8737	·8248	·7907	·7652	·7295	·6882	·6386	·5761
14	1·0909	·9370	·8581	·8082	·7732	·7471	·7103	·6675	·6159	·5500
15	1·0807	·9249	·8448	·7939	·7582	·7314	·6937	·6496	·5961	·5269
16	1·0719	·9144	·8331	·7814	·7450	·7177	·6791	·6339	·5786	·5064
17	1·0641	·9051	·8229	·7705	·7335	·7057	·6663	·6199	·5630	·4879
18	1·0572	·8970	·8138	·7607	·7232	·6950	·6549	·6075	·5516	·4712
19	1·0511	·8897	·8057	·7521	·7140	·6854	·6447	·5964	·5366	·4560
20	1·0457	·8831	·7985	·7443	·7058	·6768	·6355	·5864	·5253	·4421
21	1·0408	·8772	·7920	·7372	·6984	·6690	·6272	·5773	·5150	·4294
22	1·0363	·8719	·7860	·7309	·6916	·6620	·6196	·5691	·5056	·4176
23	1·0322	·8670	·7806	·7251	·6855	·6555	·6127	·5615	·4969	·4068
24	1·0285	·8626	·7757	·7197	·6799	·6496	·6064	·5545	·4890	·3967
25	1·0251	·8585	·7712	·7148	·6747	·6442	·6006	·5481	·4816	·3872
26	1·0220	·8548	·7670	·7103	·6699	·6392	·5952	·5422	·4748	·3784
27	1·0191	·8513	·7631	·7062	·6655	·6346	·5902	·5367	·4685	·3701
28	1·0164	·8481	·7595	·7023	·6614	·6303	·5856	·5316	·4626	·3624
29	1·0139	·8451	·7562	·6987	·6576	·6263	·5813	·5269	·4570	·3550
30	1·0116	·8423	·7531	·6954	·6540	·6226	·5773	·5224	·4519	·3481
60	·9784	·8025	·7086	·6477	·6028	·5687	·5189	·4574	·3746	·2352
∞	·9462	·7636	·6651	·5999	·5522	·5152	·4604	·3908	·2913	0

This table is reproduced by permission from "Statistical Methods for Research Workers" by Professor R. A. Fisher, published at 15/- by Oliver & Boyd Edinburgh, and attention is drawn to the larger collection in "Statistical Tables" by Professor R. A. Fisher and F. Yates, published at 12/6 by Oliver & Boyd, Edinburgh.

APPENDIX B

TABLE 10 3—5% (ROMAN TYPE) AND 1% (BOLD FACE TYPE) POINTS FOR THE DISTRIBUTION OF *F*

n_2	n_1 degrees of freedom (for greater mean square): 1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞	n_2
1	161	200	216	225	230	234	237	239	241	242	243	244	245	246	248	249	250	251	252	253	253	254	254	254	1
	4,052	**4,999**	**5,403**	**5,625**	**5,764**	**5,859**	**5,928**	**5,981**	**6,022**	**6,056**	**6,082**	**6,105**	**6,142**	**6,169**	**6,208**	**6,234**	**6,258**	**6,286**	**6,302**	**6,323**	**6,334**	**6,352**	**6,361**	**6,366**	
2	18.51	19.00	19.16	19.25	19.30	19.33	19.36	19.37	19.38	19.39	19.40	19.41	19.42	19.43	19.44	19.45	19.46	19.47	19.47	19.48	19.49	19.49	19.50	19.50	2
	98.49	**99.01**	**99.17**	**99.25**	**99.30**	**99.33**	**99.34**	**99.36**	**99.38**	**99.40**	**99.41**	**99.42**	**99.43**	**99.44**	**99.45**	**99.46**	**99.47**	**99.48**	**99.48**	**99.49**	**99.49**	**99.49**	**99.50**	**99.50**	
3	10.13	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.81	8.78	8.76	8.74	8.71	8.69	8.66	8.64	8.62	8.60	8.58	8.57	8.56	8.54	8.54	8.53	3
	34.12	**30.81**	**29.46**	**28.71**	**28.24**	**27.91**	**27.67**	**27.49**	**27.34**	**27.23**	**27.13**	**27.05**	**26.92**	**26.83**	**26.69**	**26.60**	**26.50**	**26.41**	**26.35**	**26.27**	**26.23**	**26.18**	**26.14**	**26.12**	
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.93	5.91	5.87	5.84	5.80	5.77	5.74	5.71	5.70	5.68	5.66	5.65	5.64	5.63	4
	21.20	**18.00**	**16.69**	**15.98**	**15.52**	**15.21**	**14.98**	**14.80**	**14.66**	**14.54**	**14.45**	**14.37**	**14.24**	**14.15**	**14.02**	**13.93**	**13.83**	**13.74**	**13.69**	**13.61**	**13.57**	**13.52**	**13.48**	**13.46**	
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74	4.70	4.68	4.64	4.60	4.56	4.53	4.50	4.46	4.44	4.42	4.40	4.38	4.37	4.36	5
	16.26	**13.27**	**12.06**	**11.39**	**10.97**	**10.67**	**10.45**	**10.27**	**10.15**	**10.05**	**9.96**	**9.89**	**9.77**	**9.68**	**9.55**	**9.47**	**9.38**	**9.29**	**9.24**	**9.17**	**9.13**	**9.07**	**9.04**	**9.02**	
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.96	3.92	3.87	3.84	3.81	3.77	3.75	3.72	3.71	3.69	3.68	3.67	6
	13.74	**10.92**	**9.78**	**9.15**	**8.75**	**8.47**	**8.26**	**8.10**	**7.98**	**7.87**	**7.79**	**7.72**	**7.60**	**7.52**	**7.39**	**7.31**	**7.23**	**7.14**	**7.09**	**7.02**	**6.99**	**6.94**	**6.90**	**6.88**	
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.63	3.60	3.57	3.52	3.49	3.44	3.41	3.38	3.34	3.32	3.29	3.28	3.25	3.24	3.23	7
	12.25	**9.55**	**8.45**	**7.85**	**7.46**	**7.19**	**7.00**	**6.84**	**6.71**	**6.62**	**6.54**	**6.47**	**6.35**	**6.27**	**6.15**	**6.07**	**5.98**	**5.90**	**5.85**	**5.78**	**5.75**	**5.70**	**5.67**	**5.65**	
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.34	3.31	3.28	3.23	3.20	3.15	3.12	3.08	3.05	3.03	3.00	2.98	2.96	2.94	2.93	8
	11.26	**8.65**	**7.59**	**7.01**	**6.63**	**6.37**	**6.19**	**6.03**	**5.91**	**5.82**	**5.74**	**5.67**	**5.56**	**5.48**	**5.36**	**5.28**	**5.20**	**5.11**	**5.06**	**5.00**	**4.96**	**4.91**	**4.88**	**4.86**	
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.13	3.10	3.07	3.02	2.98	2.93	2.90	2.86	2.82	2.80	2.77	2.76	2.73	2.72	2.71	9
	10.56	**8.02**	**6.99**	**6.42**	**6.06**	**5.80**	**5.62**	**5.47**	**5.35**	**5.26**	**5.18**	**5.11**	**5.00**	**4.92**	**4.80**	**4.73**	**4.64**	**4.56**	**4.51**	**4.45**	**4.41**	**4.36**	**4.33**	**4.31**	
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97	2.94	2.91	2.86	2.82	2.77	2.74	2.70	2.67	2.64	2.61	2.59	2.56	2.55	2.54	10
	10.04	**7.56**	**6.55**	**5.99**	**5.64**	**5.39**	**5.21**	**5.06**	**4.95**	**4.85**	**4.78**	**4.71**	**4.60**	**4.52**	**4.41**	**4.33**	**4.25**	**4.17**	**4.12**	**4.05**	**4.01**	**3.96**	**3.93**	**3.91**	
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.86	2.82	2.79	2.74	2.70	2.65	2.61	2.57	2.53	2.50	2.47	2.45	2.42	2.41	2.40	11
	9.65	**7.20**	**6.22**	**5.67**	**5.32**	**5.07**	**4.88**	**4.74**	**4.63**	**4.54**	**4.46**	**4.40**	**4.29**	**4.21**	**4.10**	**4.02**	**3.94**	**3.86**	**3.80**	**3.74**	**3.70**	**3.66**	**3.62**	**3.60**	
12	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80	2.76	2.72	2.69	2.64	2.60	2.54	2.50	2.46	2.42	2.40	2.36	2.35	2.32	2.31	2.30	12
	9.33	**6.93**	**5.95**	**5.41**	**5.06**	**4.82**	**4.65**	**4.50**	**4.39**	**4.30**	**4.22**	**4.16**	**4.05**	**3.98**	**3.86**	**3.78**	**3.70**	**3.61**	**3.56**	**3.49**	**3.46**	**3.41**	**3.38**	**3.36**	
13	4.67	3.80	3.41	3.18	3.02	2.92	2.84	2.77	2.72	2.67	2.63	2.60	2.55	2.51	2.46	2.42	2.38	2.34	2.32	2.28	2.26	2.24	2.22	2.21	13
	9.07	**6.70**	**5.74**	**5.20**	**4.86**	**4.62**	**4.44**	**4.30**	**4.19**	**4.10**	**4.02**	**3.96**	**3.85**	**3.78**	**3.67**	**3.59**	**3.51**	**3.42**	**3.37**	**3.30**	**3.27**	**3.21**	**3.18**	**3.16**	

This table is reproduced by permission from "Statistical Methods" by Professor George W. Snedecor, published by the Collegiate Press, Inc., Ames, Iowa.

TABLE 10 3—5% (ROMAN TYPE) AND 1% (BOLD FACE TYPE) POINTS FOR THE DISTRIBUTION OF F

n_2	n_1 degrees of freedom (for greater mean square) 1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞	n_2
14	4.60 **8.86**	3.74 **6.51**	3.34 **5.56**	3.11 **5.03**	2.96 **4.69**	2.85 **4.46**	2.77 **4.28**	2.70 **4.14**	2.65 **4.03**	2.60 **3.94**	2.56 **3.86**	2.53 **3.80**	2.48 **3.70**	2.44 **3.62**	2.39 **3.51**	2.35 **3.43**	2.31 **3.34**	2.27 **3.26**	2.24 **3.21**	2.21 **3.14**	2.19 **3.11**	2.16 **3.06**	2.14 **3.02**	2.13 **3.00**	14
15	4.54 **8.68**	3.68 **6.36**	3.29 **5.42**	3.06 **4.89**	2.90 **4.56**	2.79 **4.32**	2.70 **4.14**	2.64 **4.00**	2.59 **3.89**	2.55 **3.80**	2.51 **3.73**	2.48 **3.67**	2.43 **3.56**	2.39 **3.48**	2.33 **3.36**	2.29 **3.29**	2.25 **3.20**	2.21 **3.12**	2.18 **3.07**	2.15 **3.00**	2.12 **2.97**	2.10 **2.92**	2.08 **2.89**	2.07 **2.87**	15
16	4.49 **8.53**	3.63 **6.23**	3.24 **5.29**	3.01 **4.77**	2.85 **4.44**	2.74 **4.20**	2.66 **4.03**	2.59 **3.89**	2.54 **3.78**	2.49 **3.69**	2.45 **3.61**	2.42 **3.55**	2.37 **3.45**	2.33 **3.37**	2.28 **3.25**	2.24 **3.18**	2.20 **3.10**	2.16 **3.01**	2.13 **2.96**	2.09 **2.89**	2.07 **2.86**	2.04 **2.80**	2.02 **2.77**	2.01 **2.75**	16
17	4.45 **8.40**	3.59 **6.11**	3.20 **5.18**	2.96 **4.67**	2.81 **4.34**	2.70 **4.10**	2.62 **3.93**	2.55 **3.79**	2.50 **3.68**	2.45 **3.59**	2.41 **3.52**	2.38 **3.45**	2.33 **3.35**	2.29 **3.27**	2.23 **3.16**	2.19 **3.08**	2.15 **3.00**	2.11 **2.92**	2.08 **2.86**	2.04 **2.79**	2.02 **2.76**	1.99 **2.70**	1.97 **2.67**	1.96 **2.65**	17
18	4.41 **8.28**	3.55 **6.01**	3.16 **5.09**	2.93 **4.58**	2.77 **4.25**	2.66 **4.01**	2.58 **3.85**	2.51 **3.71**	2.46 **3.60**	2.41 **3.51**	2.37 **3.44**	2.34 **3.37**	2.29 **3.27**	2.25 **3.19**	2.19 **3.07**	2.15 **3.00**	2.11 **2.91**	2.07 **2.83**	2.04 **2.78**	2.00 **2.71**	1.98 **2.68**	1.95 **2.62**	1.93 **2.59**	1.92 **2.57**	18
19	4.38 **8.18**	3.52 **5.93**	3.13 **5.01**	2.90 **4.50**	2.74 **4.17**	2.63 **3.94**	2.55 **3.77**	2.48 **3.63**	2.43 **3.52**	2.38 **3.43**	2.34 **3.36**	2.31 **3.30**	2.26 **3.19**	2.21 **3.12**	2.15 **3.00**	2.11 **2.92**	2.07 **2.84**	2.02 **2.76**	2.00 **2.70**	1.96 **2.63**	1.94 **2.60**	1.91 **2.54**	1.90 **2.51**	1.88 **2.49**	19
20	4.35 **8.10**	3.49 **5.85**	3.10 **4.94**	2.87 **4.43**	2.71 **4.10**	2.60 **3.87**	2.52 **3.71**	2.45 **3.56**	2.40 **3.45**	2.35 **3.37**	2.31 **3.30**	2.28 **3.23**	2.23 **3.13**	2.18 **3.05**	2.12 **2.94**	2.08 **2.86**	2.04 **2.77**	1.99 **2.69**	1.96 **2.63**	1.92 **2.56**	1.90 **2.53**	1.87 **2.47**	1.85 **2.44**	1.84 **2.42**	20
21	4.32 **8.02**	3.47 **5.78**	3.07 **4.87**	2.84 **4.37**	2.68 **4.04**	2.57 **3.81**	2.49 **3.65**	2.42 **3.51**	2.37 **3.40**	2.32 **3.31**	2.28 **3.24**	2.25 **3.17**	2.20 **3.07**	2.15 **2.99**	2.09 **2.88**	2.05 **2.80**	2.00 **2.72**	1.96 **2.63**	1.93 **2.58**	1.89 **2.51**	1.87 **2.47**	1.84 **2.42**	1.82 **2.38**	1.81 **2.36**	21
22	4.30 **7.94**	3.44 **5.72**	3.05 **4.82**	2.82 **4.31**	2.66 **3.99**	2.55 **3.76**	2.47 **3.59**	2.40 **3.45**	2.35 **3.35**	2.30 **3.26**	2.26 **3.18**	2.23 **3.12**	2.18 **3.02**	2.13 **2.94**	2.07 **2.83**	2.03 **2.75**	1.98 **2.67**	1.93 **2.58**	1.91 **2.53**	1.87 **2.46**	1.84 **2.42**	1.81 **2.37**	1.80 **2.33**	1.78 **2.31**	22
23	4.28 **7.88**	3.42 **5.66**	3.03 **4.76**	2.80 **4.26**	2.64 **3.94**	2.53 **3.71**	2.45 **3.54**	2.38 **3.41**	2.32 **3.30**	2.28 **3.21**	2.24 **3.14**	2.20 **3.07**	2.14 **2.97**	2.10 **2.89**	2.04 **2.78**	2.00 **2.70**	1.96 **2.62**	1.91 **2.53**	1.88 **2.48**	1.84 **2.41**	1.82 **2.37**	1.79 **2.32**	1.77 **2.28**	1.76 **2.26**	23
24	4.26 **7.82**	3.40 **5.61**	3.01 **4.72**	2.78 **4.22**	2.62 **3.90**	2.51 **3.67**	2.43 **3.50**	2.36 **3.36**	2.30 **3.25**	2.26 **3.17**	2.22 **3.09**	2.18 **3.03**	2.13 **2.93**	2.09 **2.85**	2.02 **2.74**	1.98 **2.66**	1.94 **2.58**	1.89 **2.49**	1.86 **2.44**	1.82 **2.36**	1.80 **2.33**	1.76 **2.27**	1.74 **2.23**	1.73 **2.21**	24
25	4.24 **7.77**	3.38 **5.57**	2.99 **4.68**	2.76 **4.18**	2.60 **3.86**	2.49 **3.63**	2.41 **3.46**	2.34 **3.32**	2.28 **3.21**	2.24 **3.13**	2.20 **3.05**	2.16 **2.99**	2.11 **2.89**	2.06 **2.81**	2.00 **2.70**	1.96 **2.62**	1.92 **2.54**	1.87 **2.45**	1.84 **2.40**	1.80 **2.32**	1.77 **2.29**	1.74 **2.23**	1.72 **2.19**	1.71 **2.17**	25
26	4.22 **7.72**	3.37 **5.53**	2.98 **4.64**	2.74 **4.14**	2.59 **3.82**	2.47 **3.59**	2.39 **3.42**	2.32 **3.29**	2.27 **3.17**	2.22 **3.09**	2.18 **3.02**	2.15 **2.96**	2.10 **2.86**	2.05 **2.77**	1.99 **2.66**	1.95 **2.58**	1.90 **2.50**	1.85 **2.41**	1.82 **2.36**	1.78 **2.28**	1.76 **2.25**	1.72 **2.19**	1.70 **2.15**	1.69 **2.13**	26

The function, $F = e$ with exponent $2z$, is computed in part from Fisher's table VI (7). Additional entries are by interpolation, mostly graphical.

This table is reproduced by permission from "Statistical Methods" by Professor George W. Snedecor, published by the Collegiate Press, Inc., Ames, Iowa.

TABLE 10.3—5% (ROMAN TYPE) AND 1% (BOLD FACE TYPE) POINTS FOR THE DISTRIBUTION OF *F*

n_2	n_1 degrees of freedom (for greater mean square)																								n_2
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞	
27	4.21 **7.68**	3.35 **5.49**	2.96 **4.60**	2.73 **4.11**	2.57 **3.79**	2.46 **3.56**	2.37 **3.39**	2.30 **3.26**	2.25 **3.14**	2.20 **3.06**	2.16 **2.98**	2.13 **2.93**	2.08 **2.83**	2.03 **2.74**	1.97 **2.63**	1.93 **2.55**	1.88 **2.47**	1.84 **2.38**	1.80 **2.33**	1.76 **2.25**	1.74 **2.21**	1.71 **2.16**	1.68 **2.12**	1.67 **2.10**	27
28	4.20 **7.64**	3.34 **5.45**	2.95 **4.57**	2.71 **4.07**	2.56 **3.76**	2.44 **3.53**	2.36 **3.36**	2.29 **3.23**	2.24 **3.11**	2.19 **3.03**	2.15 **2.95**	2.12 **2.90**	2.06 **2.80**	2.02 **2.71**	1.96 **2.60**	1.91 **2.52**	1.87 **2.44**	1.81 **2.35**	1.78 **2.30**	1.75 **2.22**	1.72 **2.18**	1.69 **2.13**	1.67 **2.09**	1.65 **2.06**	28
29	4.18 **7.60**	3.33 **5.42**	2.93 **4.54**	2.70 **4.04**	2.54 **3.73**	2.43 **3.50**	2.35 **3.33**	2.28 **3.20**	2.22 **3.08**	2.18 **3.00**	2.14 **2.92**	2.10 **2.87**	2.05 **2.77**	2.00 **2.68**	1.94 **2.57**	1.90 **2.49**	1.85 **2.41**	1.80 **2.32**	1.77 **2.27**	1.73 **2.19**	1.71 **2.15**	1.68 **2.10**	1.65 **2.06**	1.64 **2.03**	29
30	4.17 **7.56**	3.32 **5.39**	2.92 **4.51**	2.69 **4.02**	2.53 **3.70**	2.42 **3.47**	2.34 **3.30**	2.27 **3.17**	2.21 **3.06**	2.16 **2.98**	2.12 **2.90**	2.09 **2.84**	2.04 **2.74**	1.99 **2.66**	1.93 **2.55**	1.89 **2.47**	1.84 **2.38**	1.79 **2.29**	1.76 **2.24**	1.72 **2.16**	1.69 **2.13**	1.66 **2.07**	1.64 **2.03**	1.62 **2.01**	30
32	4.15 **7.50**	3.30 **5.34**	2.90 **4.46**	2.67 **3.97**	2.51 **3.66**	2.40 **3.42**	2.32 **3.25**	2.25 **3.12**	2.19 **3.01**	2.14 **2.94**	2.10 **2.86**	2.07 **2.80**	2.02 **2.70**	1.97 **2.62**	1.91 **2.51**	1.86 **2.42**	1.82 **2.34**	1.76 **2.25**	1.74 **2.20**	1.69 **2.12**	1.67 **2.08**	1.64 **2.02**	1.61 **1.98**	1.59 **1.96**	32
34	4.13 **7.44**	3.28 **5.29**	2.88 **4.42**	2.65 **3.93**	2.49 **3.61**	2.38 **3.38**	2.30 **3.21**	2.23 **3.08**	2.17 **2.97**	2.12 **2.89**	2.08 **2.82**	2.05 **2.76**	2.00 **2.66**	1.95 **2.58**	1.89 **2.47**	1.84 **2.38**	1.80 **2.30**	1.74 **2.21**	1.71 **2.15**	1.67 **2.08**	1.64 **2.04**	1.61 **1.98**	1.59 **1.94**	1.57 **1.91**	34
36	4.11 **7.39**	3.26 **5.25**	2.86 **4.38**	2.63 **3.89**	2.48 **3.58**	2.36 **3.35**	2.28 **3.18**	2.21 **3.04**	2.15 **2.94**	2.10 **2.86**	2.06 **2.78**	2.03 **2.72**	1.98 **2.62**	1.93 **2.54**	1.87 **2.43**	1.82 **2.35**	1.78 **2.26**	1.72 **2.17**	1.69 **2.12**	1.65 **2.04**	1.62 **2.00**	1.59 **1.94**	1.56 **1.90**	1.55 **1.87**	36
38	4.10 **7.35**	3.25 **5.21**	2.85 **4.34**	2.62 **3.86**	2.46 **3.54**	2.35 **3.32**	2.26 **3.15**	2.19 **3.02**	2.14 **2.91**	2.09 **2.82**	2.05 **2.75**	2.02 **2.69**	1.96 **2.59**	1.92 **2.51**	1.85 **2.40**	1.80 **2.32**	1.76 **2.22**	1.71 **2.14**	1.67 **2.08**	1.63 **2.00**	1.60 **1.97**	1.57 **1.90**	1.54 **1.86**	1.53 **1.84**	38
40	4.08 **7.31**	3.23 **5.18**	2.84 **4.31**	2.61 **3.83**	2.45 **3.51**	2.34 **3.29**	2.25 **3.12**	2.18 **2.99**	2.12 **2.88**	2.07 **2.80**	2.04 **2.73**	2.00 **2.66**	1.95 **2.56**	1.90 **2.49**	1.84 **2.37**	1.79 **2.29**	1.74 **2.20**	1.69 **2.11**	1.66 **2.05**	1.61 **1.97**	1.59 **1.94**	1.55 **1.88**	1.53 **1.84**	1.51 **1.81**	40
42	4.07 **7.27**	3.22 **5.15**	2.83 **4.29**	2.59 **3.80**	2.44 **3.49**	2.32 **3.26**	2.24 **3.10**	2.17 **2.96**	2.11 **2.86**	2.06 **2.77**	2.02 **2.70**	1.99 **2.64**	1.94 **2.54**	1.89 **2.46**	1.82 **2.35**	1.78 **2.26**	1.73 **2.17**	1.68 **2.08**	1.64 **2.02**	1.60 **1.94**	1.57 **1.91**	1.54 **1.85**	1.51 **1.80**	1.49 **1.78**	42
44	4.06 **7.24**	3.21 **5.12**	2.82 **4.26**	2.58 **3.78**	2.43 **3.46**	2.31 **3.24**	2.23 **3.07**	2.16 **2.94**	2.10 **2.84**	2.05 **2.75**	2.01 **2.68**	1.98 **2.62**	1.92 **2.52**	1.88 **2.44**	1.81 **2.32**	1.76 **2.24**	1.72 **2.15**	1.66 **2.06**	1.63 **2.00**	1.58 **1.92**	1.56 **1.88**	1.52 **1.82**	1.50 **1.78**	1.48 **1.75**	44
46	4.05 **7.21**	3.20 **5.10**	2.81 **4.24**	2.57 **3.76**	2.42 **3.44**	2.30 **3.22**	2.22 **3.05**	2.14 **2.92**	2.09 **2.82**	2.04 **2.73**	2.00 **2.66**	1.97 **2.60**	1.91 **2.50**	1.87 **2.42**	1.80 **2.30**	1.75 **2.22**	1.71 **2.13**	1.65 **2.04**	1.62 **1.98**	1.57 **1.90**	1.54 **1.86**	1.51 **1.80**	1.48 **1.76**	1.46 **1.72**	46
48	4.04 **7.19**	3.19 **5.08**	2.80 **4.22**	2.56 **3.74**	2.41 **3.42**	2.30 **3.20**	2.21 **3.04**	2.14 **2.90**	2.08 **2.80**	2.03 **2.71**	1.99 **2.64**	1.96 **2.58**	1.90 **2.48**	1.86 **2.40**	1.79 **2.28**	1.74 **2.20**	1.70 **2.11**	1.64 **2.02**	1.61 **1.96**	1.56 **1.88**	1.53 **1.84**	1.50 **1.78**	1.47 **1.73**	1.45 **1.70**	48

This table is reproduced by permission from "Statistical Methods" by Professor George W. Snedecor, published by the Collegiate Press, Inc., Ames, Iowa.

TABLE 10.3—5% (ROMAN TYPE) AND 1% (BOLD FACE TYPE) POINTS FOR THE DISTRIBUTION OF F

n_1 degrees of freedom (for greater mean square)

n_2	1	2	3	4	5	6	7	8	n_2
50	4.03 **7.17**	3.18 **5.06**	2.79 **4.20**	2.56 **3.72**	2.40 **3.41**	2.29 **3.18**	2.20 **3.02**	2.13 **2.88**	50
55	4.02 **7.12**	3.17 **5.01**	2.78 **4.16**	2.54 **3.68**	2.38 **3.37**	2.27 **3.15**	2.18 **2.98**	2.11 **2.85**	55
60	4.00 **7.08**	3.15 **4.98**	2.76 **4.13**	2.52 **3.65**	2.37 **3.34**	2.25 **3.12**	2.17 **2.95**	2.10 **2.82**	60
65	3.99 **7.04**	3.14 **4.95**	2.75 **4.10**	2.51 **3.62**	2.36 **3.31**	2.24 **3.09**	2.15 **2.93**	2.08 **2.79**	65
70	3.98 **7.01**	3.13 **4.92**	2.74 **4.08**	2.50 **3.60**	2.35 **3.29**	2.23 **3.07**	2.14 **2.91**	2.07 **2.77**	70
80	3.96 **6.96**	3.11 **4.88**	2.72 **4.04**	2.48 **3.56**	2.33 **3.25**	2.21 **3.04**	2.12 **2.87**	2.05 **2.74**	80
100	3.94 **6.90**	3.09 **4.82**	2.70 **3.98**	2.46 **3.51**	2.30 **3.20**	2.19 **2.99**	2.10 **2.82**	2.03 **2.69**	100
125	3.92 **6.84**	3.07 **4.78**	2.68 **3.94**	2.44 **3.47**	2.29 **3.17**	2.17 **2.95**	2.08 **2.79**	2.01 **2.65**	125
150	3.91 **6.81**	3.06 **4.75**	2.67 **3.91**	2.43 **3.44**	2.27 **3.14**	2.16 **2.92**	2.07 **2.76**	2.00 **2.62**	150
200	3.89 **6.76**	3.04 **4.71**	2.65 **3.88**	2.41 **3.41**	2.26 **3.11**	2.14 **2.90**	2.05 **2.73**	1.98 **2.60**	200
400	3.86 **6.70**	3.02 **4.66**	2.62 **3.83**	2.39 **3.36**	2.23 **3.06**	2.12 **2.85**	2.03 **2.69**	1.96 **2.55**	400
1000	3.85 **6.66**	3.00 **4.62**	2.61 **3.80**	2.38 **3.34**	2.22 **3.04**	2.10 **2.82**	2.02 **2.66**	1.95 **2.53**	1000
∞	3.84 **6.64**	2.99 **4.60**	2.60 **3.78**	2.37 **3.32**	2.21 **3.02**	2.09 **2.80**	2.01 **2.64**	1.94 **2.51**	∞

n_2	9	10	11	12	14	16	20	24	n_2
50	2.07 **2.78**	2.03 **2.70**	1.98 **2.62**	1.95 **2.56**	1.90 **2.46**	1.85 **2.39**	1.78 **2.26**	1.74 **2.18**	50
55	2.05 **2.75**	2.00 **2.66**	1.97 **2.59**	1.93 **2.53**	1.88 **2.43**	1.83 **2.35**	1.76 **2.23**	1.72 **2.15**	55
60	2.04 **2.72**	1.99 **2.63**	1.95 **2.56**	1.92 **2.50**	1.86 **2.40**	1.81 **2.32**	1.75 **2.20**	1.70 **2.12**	60
65	2.02 **2.70**	1.98 **2.61**	1.94 **2.54**	1.90 **2.47**	1.85 **2.37**	1.80 **2.30**	1.73 **2.18**	1.68 **2.09**	65
70	2.01 **2.67**	1.97 **2.59**	1.93 **2.51**	1.89 **2.45**	1.84 **2.35**	1.79 **2.28**	1.72 **2.15**	1.67 **2.07**	70
80	1.99 **2.64**	1.95 **2.55**	1.91 **2.48**	1.88 **2.41**	1.82 **2.32**	1.77 **2.24**	1.70 **2.11**	1.65 **2.03**	80
100	1.97 **2.59**	1.92 **2.51**	1.88 **2.43**	1.85 **2.36**	1.79 **2.26**	1.75 **2.19**	1.68 **2.06**	1.63 **1.98**	100
125	1.95 **2.56**	1.90 **2.47**	1.86 **2.40**	1.83 **2.33**	1.77 **2.23**	1.72 **2.15**	1.65 **2.03**	1.60 **1.94**	125
150	1.94 **2.53**	1.89 **2.44**	1.85 **2.37**	1.82 **2.30**	1.76 **2.20**	1.71 **2.12**	1.64 **2.00**	1.59 **1.91**	150
200	1.92 **2.50**	1.87 **2.41**	1.83 **2.34**	1.80 **2.28**	1.74 **2.17**	1.69 **2.09**	1.62 **1.97**	1.57 **1.88**	200
400	1.90 **2.46**	1.85 **2.37**	1.81 **2.29**	1.78 **2.23**	1.72 **2.12**	1.67 **2.04**	1.60 **1.92**	1.54 **1.84**	400
1000	1.89 **2.43**	1.84 **2.34**	1.80 **2.26**	1.76 **2.20**	1.70 **2.09**	1.65 **2.01**	1.58 **1.89**	1.53 **1.81**	1000
∞	1.88 **2.41**	1.83 **2.32**	1.79 **2.24**	1.75 **2.18**	1.69 **2.07**	1.64 **1.99**	1.57 **1.87**	1.52 **1.79**	∞

n_2	30	40	50	75	100	200	500	∞	n_2
50	1.69 **2.10**	1.63 **2.00**	1.60 **1.94**	1.55 **1.86**	1.52 **1.82**	1.48 **1.76**	1.46 **1.71**	1.44 **1.68**	50
55	1.67 **2.06**	1.61 **1.96**	1.58 **1.90**	1.52 **1.82**	1.50 **1.78**	1.46 **1.71**	1.43 **1.66**	1.41 **1.64**	55
60	1.65 **2.03**	1.59 **1.93**	1.56 **1.87**	1.50 **1.79**	1.48 **1.74**	1.44 **1.68**	1.41 **1.63**	1.39 **1.60**	60
65	1.63 **2.00**	1.57 **1.90**	1.54 **1.84**	1.49 **1.76**	1.46 **1.71**	1.42 **1.64**	1.39 **1.60**	1.37 **1.56**	65
70	1.62 **1.98**	1.56 **1.88**	1.53 **1.82**	1.47 **1.74**	1.45 **1.69**	1.40 **1.62**	1.37 **1.56**	1.35 **1.53**	70
80	1.60 **1.94**	1.54 **1.84**	1.51 **1.78**	1.45 **1.70**	1.42 **1.65**	1.38 **1.57**	1.35 **1.52**	1.32 **1.49**	80
100	1.57 **1.89**	1.51 **1.79**	1.48 **1.73**	1.42 **1.64**	1.39 **1.59**	1.34 **1.51**	1.30 **1.46**	1.28 **1.43**	100
125	1.55 **1.85**	1.49 **1.75**	1.45 **1.68**	1.39 **1.59**	1.36 **1.54**	1.31 **1.46**	1.27 **1.40**	1.25 **1.37**	125
150	1.54 **1.83**	1.47 **1.72**	1.44 **1.66**	1.37 **1.56**	1.34 **1.51**	1.29 **1.43**	1.25 **1.37**	1.22 **1.33**	150
200	1.52 **1.79**	1.45 **1.69**	1.42 **1.62**	1.35 **1.53**	1.32 **1.48**	1.26 **1.39**	1.22 **1.33**	1.19 **1.28**	200
400	1.49 **1.74**	1.42 **1.64**	1.38 **1.57**	1.32 **1.47**	1.28 **1.42**	1.22 **1.32**	1.16 **1.24**	1.13 **1.19**	400
1000	1.47 **1.71**	1.41 **1.61**	1.36 **1.54**	1.30 **1.44**	1.26 **1.38**	1.19 **1.28**	1.13 **1.19**	1.08 **1.11**	1000
∞	1.46 **1.69**	1.40 **1.59**	1.35 **1.52**	1.28 **1.41**	1.24 **1.36**	1.17 **1.25**	1.11 **1.15**	1.00 **1.00**	∞

APPENDIX C

TABLE IV. *5 per cent. limits for L_1.*

f	2	3	4	5	6	7	8	9	11	14	19	29	59	∞
k \ n	3	4	5	6	7	8	9	10	12	15	20	30	60	∞
2	·312	·478	·585	·656	·708	·745	·775	·798	·833	·868	·902	·935	·968	1·000
3	·304	·470	·576	·648	·700	·739	·769	·792	·828	·863	·898	·933	·967	1·000
4	·315	·480	·585	·656	·707	·744	·774	·797	·832	·866	·900	·934	·967	1·000
5	·328	·491	·595	·665	·714	·751	·780	·802	·836	·870	·903	·936	·968	1·000
6	·339	·502	·604	·673	·721	·757	·785	·808	·841	·873	·906	·938	·969	1·000
7	·350	·512	·612	·680	·727	·763	·790	·812	·844	·876	·908	·939	·970	1·000
8	·359	·520	·620	·686	·733	·768	·795	·816	·848	·879	·910	·941	·971	1·000
9	·367	·527	·626	·691	·738	·772	·798	·819	·851	·881	·912	·942	·971	1·000
10	·374	·534	·631	·696	·742	·776	·802	·822	·853	·883	·913	·943	·972	1·000
12	·387	·545	·641	·704	·749	·782	·807	·828	·857	·887	·916	·944	·973	1·000
14	·397	·554	·649	·711	·755	·787	·812	·832	·861	·890	·918	·946	·973	1·000
16	·405	·561	·655	·716	·759	·791	·816	·835	·863	·892	·920	·947	·974	1·000
18	·412	·567	·660	·721	·763	·795	·819	·838	·866	·894	·921	·948	·974	1·000
20	·418	·573	·665	·725	·767	·798	·822	·840	·868	·896	·922	·949	·975	1·000
22	·424	·577	·669	·728	·770	·800	·824	·843	·870	·897	·924	·950	·975	1·000
24	·428	·581	·672	·731	·772	·802	·826	·844	·872	·898	·924	·950	·975	1·000
26	·433	·585	·675	·734	·775	·805	·828	·846	·873	·899	·925	·951	·976	1·000
28	·437	·589	·678	·736	·777	·807	·829	·848	·874	·900	·926	·951	·976	1·000
30	·441	·592	·681	·739	·779	·809	·831	·849	·876	·901	·927	·952	·976	1·000

TABLE V. *1 per cent. limits for L_1.*

f	2	3	4	5	6	7	8	9	11	14	19	29	59	∞
k \ n	3	4	5	6	7	8	9	10	12	15	20	30	60	∞
2	·141	·284	·398	·485	·551	·603	·645	·678	·730	·783	·836	·890	·945	1·000
3	·162	·314	·429	·514	·578	·628	·667	·699	·748	·798	·848	·898	·949	1·000
4	·188	·345	·459	·542	·604	·652	·689	·719	·765	·812	·859	·906	·953	1·000
5	·210	·370	·484	·565	·624	·670	·706	·735	·779	·823	·867	·911	·956	1·000
6	·230	·391	·504	·583	·641	·685	·720	·748	·789	·832	·874	·916	·958	1·000
7	·246	·409	·520	·597	·654	·697	·730	·757	·798	·839	·879	·920	·960	1·000
8	·260	·424	·534	·610	·665	·707	·740	·766	·805	·844	·884	·923	·962	1·000
9	·273	·437	·545	·620	·674	·715	·747	·773	·811	·849	·887	·925	·963	1·000
10	·284	·448	·555	·629	·682	·722	·753	·779	·816	·853	·890	·927	·964	1·000
12	·303	·467	·572	·644	·696	·734	·764	·789	·824	·860	·896	·931	·966	1·000
14	·318	·481	·585	·655	·706	·744	·773	·796	·831	·865	·900	·933	·967	1·000
16	·331	·493	·596	·665	·714	·751	·779	·802	·836	·870	·903	·936	·968	1·000
18	·342	·504	·605	·672	·721	·756	·784	·807	·840	·873	·905	·937	·969	1·000
20	·352	·512	·613	·679	·727	·761	·788	·811	·844	·876	·908	·939	·970	1·000
22	·360	·520	·619	·684	·732	·765	·792	·814	·847	·878	·909	·940	·970	1·000
24	·367	·526	·624	·688	·736	·768	·795	·817	·850	·880	·911	·941	·971	1·000
26	·373	·532	·629	·693	·740	·772	·798	·820	·852	·882	·912	·942	·971	1·000
28	·379	·537	·634	·697	·744	·776	·802	·823	·854	·884	·914	·943	·972	1·000
30	·386	·543	·639	·703	·748	·781	·806	·827	·856	·886	·915	·944	·972	1·000

These tables are reproduced by permission from "Statistical Research Memoirs", Volume I, edited by J. Neyman and E. S. Pearson and issued by the Department of Statistics, University of London, University College, London, England.